

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

**Computação Afectiva –  
implementação de um *wearable* multimédia**

Armindo José Gonçalves Dias

Licenciado em Engenharia Informática e Computação  
pela Universidade de Engenharia da Universidade do Porto

Dissertação submetida para satisfação parcial dos  
requisitos do grau de mestre  
em  
Tecnologia Multimédia

Dissertação realizada sob a supervisão de  
Professor Eurico Carrapatoso,  
do Departamento de Engenharia Electrotécnica e de Computadores  
da Faculdade de Engenharia da Universidade do Porto  
e  
Professor André Puga,  
do Departamento de Engenharia Mecânica e Gestão Industrial  
da Faculdade de Engenharia da Universidade do Porto

Porto, Janeiro de 2005

## Resumo

A Computação Afectiva, embora se trate ainda de uma área de investigação recente, pode ter um papel fundamental para o avanço na forma como a emoção e a cognição são compreendidas, resultando em melhorias na forma como os computadores são capazes de resolver problemas, em avanços nas possibilidades de interacção homem-máquina e na forma como os computadores poderão influenciar o próprio desenvolvimento humano.

Esta Dissertação de Mestrado descreve a revisão do estado-da-arte em Computação Afectiva, a análise de ferramentas de programação e de recolha e processamento de sinais fisiológicos, a análise de requisitos e arquitectura de um protótipo de Computação Afectiva, e, finalmente, a sua implementação, aplicação e teste num *wearable* construído para o efeito.

O protótipo desenvolvido capta os sinais fisiológicos de um indivíduo e utiliza-os para interagir em tempo real com conteúdos multimédia (imagem, vídeo, som e texto). Além disso, e com base nesses mesmos sinais fisiológicos, reconhece o estado emotivo de tristeza, alegria, susto e repouso do indivíduo, reflectindo-se também cada um desses estados na interacção com todo o conteúdo multimédia. Utilizaram-se para o efeito três sinais fisiológicos: o batimento cardíaco, a condutividade eléctrica da pele e a actividade muscular.

O protótipo é constituído por um módulo de aquisição dos sinais fisiológicos, um módulo de apresentação de conteúdos multimédia, um módulo de detecção do estado emotivo do utilizador através de uma implementação de redes neuronais e um módulo de histórico onde são registadas todas as ocorrências dos estados emotivos detectados ao longo da utilização do sistema.

A avaliação do funcionamento em tempo-real do protótipo quando utilizado no *wearable* apresenta resultados bastante positivos, detectando em tempo-real o estado emotivo de alegria, repouso e susto, sendo o estado da tristeza não tão facilmente detectado.

## Abstract

Affective Computing, although still a recent research area, may have a fundamental importance for the development of how emotion and cognition are understood, resulting in better problem solving capabilities of current computer systems, developments on the possibilities of human-computer interaction and in the way computers can influence the human development

This Thesis describes about the state-of-the-art in the area of Affective Computing, the analysis of programming languages and tools for the acquisition and analysis of physiological signals, the requirements analysis and architecture for a prototype of Affective Computing and, finally, its implementation and test on a wearable.

The developed prototype acquires the physiological signals of an individual and uses them to interact in real time with multimedia contents (image, video, sound and text). Besides that, and with the same physiological signals, it recognizes the emotive state of sadness, happiness, surprise and rest of an individual, and each one of these states is also reflected on the interaction with all the multimedia content. The three physiological signals used were: the heart beat, the electrical skin conductivity and the muscular activity.

The prototype is composed by a module of signal acquisition, a module for the presentation of multimedia content, a module for the detection of the emotional state of the individual through an implementation of neural networks and a logging module for the recording of the emotional states detected while the system is in use.

The evaluation of the prototype functioning in real time with the wearable gives good results, detecting the emotional state of happiness, surprise and rest in real time, being the emotional state of sadness not so easily detected.

## Agradecimentos

Gostava de em primeiro lugar agradecer ao Professor Eurico Carrapatoso pela orientação sábia, pelos conselhos sempre úteis e por toda a atenção, compreensão e dedicação nunca negadas ao longo de todo o trabalho realizado.

Gostava também de agradecer ao Professor André Puga por toda a sua disponibilidade, dedicação e apoio imprescindíveis para a elaboração deste trabalho.

Agradeço ainda à Luísa Lamas pelas palavras de incentivo sempre presentes e por toda a ajuda que nunca hesitou em me conceder.

Finalmente, mas não menos importante, agradeço à Joana Lessa, a pessoa sem a qual este trabalho nunca teria existido.

A todas as pessoas que, directa ou indirectamente, contribuíram para a elaboração deste trabalho, os meus sinceros agradecimentos.

# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Emoções e Computação afectiva</b>	<b>5</b>
2.1	Emoções: fisiologia e cogniscência	5
2.1.1	Comunicação afectiva	8
2.1.2	Aspectos cognitivos das emoções	12
2.2	Computadores mais humanos	14
2.2.1	Computadores que reconhecem emoções	15
2.2.2	Computadores que exprimem emoções	16
2.2.3	Computadores que possuem emoções	18
2.3	Aplicações da computação afectiva	24
2.3.1	<i>Wearables</i> afectivos	24
2.3.2	Outras aplicações	26
2.4	Sinais Fisiológicos	28
2.4.1	Batimento cardíaco	30
2.4.2	Condutividade eléctrica da pele	32
2.4.3	Actividade muscular	35
<b>3</b>	<b>Ferramentas</b>	<b>38</b>
3.1	PD – <i>Pure Data</i>	38
3.1.1	Introdução	39
3.1.2	GEM – <i>Graphics Environment for Multimedia</i>	40
3.1.3	Lista de Objectos do GEM	40
3.1.4	Aplicações do PD/GEM	42
3.2	Max/MSP e Jitter	43
3.3	MIDI – <i>Musical Instrument Digital Interface</i>	46
3.4	WaveRider	47
3.5	Outros sistemas de captação de sinais fisiológicos	48

3.5.1 Procomp+/Biograph	48
3.5.2 Biomuse	49
3.5.3 IBVA – <i>Interactive Brainwave Visual Analyser</i>	49
3.6 MIDI YOKE - MIDI <i>Patch Cable driver</i>	50
3.7 Java	51
3.8 Matlab	52
3.9 Redes neuronais	53
3.9.1 Modelo do neurónio artificial e arquitectura da rede neuronal	55
3.9.1.1 Neurónio com entrada simples	55
3.9.1.2 Exemplos de funções de transferência	56
3.9.1.3 Neurónio com entradas múltiplas	57
3.9.1.4 Uma camada de neurónios	58
3.9.1.5 Camadas múltiplas de neurónios	58
3.9.1.6 Arquitecturas recursivas	59
3.9.2 Aprendizagem	59
3.9.3 Métodos de treino de redes neuronais	60
3.9.3.1 Perceptron	60
3.9.3.2 Regra de Widrow-Hoff Learning	60
3.9.3.3 Backpropagation	61
3.10 PHP	61
<b>4 Definição do sistema e especificação</b>	<b>64</b>
4.1 Parte I – Visão geral do sistema	64
4.1.1 Utilização do sistema	64
4.1.2 Arquitectura e interacção com outros sistemas	65
4.2 Parte II – Requisitos Funcionais	66
4.2.1 Descrição das funcionalidades	66
4.2.2 Resposta do sistema	69
4.2.3 Calendarização da implementação do sistema	69
4.3 Arquitectura do sistema	70

<b>5</b>	<b>Detalhes de implementação e resultados</b>	<b>72</b>
5.1	Arquitectura do Sistema	72
5.2	Módulo de Apresentação	74
5.2.1	Pacotes	75
5.2.2	Distribuição em rede do sistema	87
5.3	Módulo de Aquisição dos Sinais Fisiológicos	88
5.3.1	WaveRider Jr. e WaveWare	89
5.3.2	Interligação entre o Módulo de Apresentação e o WaveRider	91
5.4	Módulo de Reconhecimento de Emoções	94
5.4.1	Recolha de dados fisiológicos	95
5.4.2	Tratamento dos dados em Matlab e redes neuronais	97
5.4.3	Implementação em tempo-real do sistema de reconhecimento de emoções	103
5.5	Módulo do Histórico do Sistema	110
5.5.1	Registo de informação	111
5.5.2	Leitura de informação	114
5.6	<i>Wearable</i> afectivo	115
5.7	Inquéritos de opinião	116
5.8	Resultados de utilização do sistema	118
<b>6</b>	<b>Conclusão</b>	<b>120</b>
6.1	Breve retrospectiva do trabalho realizado	120
6.2	Desenvolvimentos futuros	121
6.3	Considerações finais	124
	<b>Referências</b>	<b>126</b>

## Lista de Figuras

Figura 2.1 – Critérios utilizados no reconhecimento emocional por parte do computador	16
Figura 2.2 – Critérios de expressividade de emoções	18
Figura 2.3 – Componentes de um sistema emocional	21
Figura 2.4 – Etapas do reconhecimento das emoções	29
Figura 2.5 – Fases do batimento cardíaco	31
Figura 2.6 – Ondas P, Q, R, S e T do batimento cardíaco	32
Figura 2.7 – Exemplo de um registo de ECG	32
Figura 2.8 – Exemplo de um registo de GSR	35
Figura 2.9 – Exemplo de um registo de EMG	37
Figura 3.1 – Exemplo de dois <i>patches</i> em PD	39
Figura 3.2 – Max	44
Figura 3.3 – MSP	45
Figura 3.4 – Jitter	45
Figura 3.5 – WaveWare	48
Figura 3.6 – ProComp +	48
Figura 3.7 – BioMuse	49
Figura 3.8 – Sistema IBVA	50
Figura 3.9 – Representação de um neurónio artificial	55
Figura 3.10 – Função de transferência Hard Limit	56
Figura 3.11 – Função de transferência Linear	56
Figura 3.12 – Função de transferência Sigmoid	57
Figura 3.13 – Representação de um neurónio com R entradas	57
Figura 3.14 – Representação de uma camada de neurónios com R entradas e S neurónios	58
Figura 3.15 – Representação de uma arquitectura neuronal recursiva com S neurónios	59
Figura 4.1 – Interligação dos módulos constituintes do sistema	71
Figura 5.1 – Detalhe da interligação dos Módulos constituintes do sistema	73
Figura 5.2 – Caixa de informação com excerto de um poema no Pacote do Repouso	75



Figura 5.3 – Conjunto de portas presentes no Pacote de Entrada	76
Figura 5.4 – Sequência de cortinas mostradas na transição para um novo estado emotivo	76
Figura 5.5 – <i>Patch</i> remote_ogg_player.pd	78
Figura 5.6 – Nível 1 do Pacote da Tristeza	79
Figura 5.7 – Nível 2 do Pacote da Tristeza	80
Figura 5.8 – Nível 3 do Pacote da Tristeza	81
Figura 5.9 – Alternância das tiras de imagens constituintes do Pacote do Susto	82
Figura 5.10 – FM7, Sintetizador Virtual VST da Native Instruments	84
Figura 5.11 – <i>Patch</i> Scale.pd responsável pela criação de uma melodia MIDI em determinada escala	84
Figura 5.12 – <i>Patch</i> remote_vst_player.pd para a inclusão de instrumentos VST no PD	85
Figura 5.13 – Sequência de entrada no Pacote da Alegria	85
Figura 5.14 – Pacote do Repouso	87
Figura 5.15 – Utilização de [netsend] e [netreceive] para a comunicação entre <i>patches</i> distribuídos por 2 computadores	88
Figura 5.16 – Colocação correcta dos eléctrodos para o sinal GSR, Hr e muscular	89
Figura 5.17 – Opções do WaveWare para a configuração do sinal GSR, HR e muscular	90
Figura 5.18 – WaveWare a captar em simultâneo o sinal HR, GSR e muscular	90
Figura 5.19 – Caixa de diálogo do WaveWare para a selecção do dispositivo MIDI de saída	92
Figura 5.20 – Caixa de diálogo da saída MIDI para um sinal, neste caso o sinal GSR no canal MIDI 1	92
Figura 5.21 – Opções de transmissão MIDI para o sinal captado, neste caso, canal MIDI 2 para o HR	93
Figura 5.22 – Comando 'pd -listdev' onde é possível visualizar as entradas e saídas MIDI disponibilizadas no PD	94
Figura 5.23 – Janela do Matlab da saída da rede neuronal, detectando o estado 3, alegria	101
Figura 5.24 – Interligação e passagem de informação entre <i>packages</i> de desenvolvimento diferentes. Arquitectura cliente-servidor ou <i>three-tier</i>	104
Figura 5.25 – <i>Patch</i> mat_brain.pd	105
Figura 5.26 - Parte do <i>patch</i> brain.pd responsável pela recepção e escrita dos dados MIDI enviados pelo WaveWare	106

Figura 5.27 – Parte do <i>patch</i> brain.pd responsável pela análise feita em Java e Matlab. O resultado obtido dessa análise é recebido no primeiro <i>outlet</i> , <i>received data</i>	107
Figura 5.28 – <i>Patch</i> sub_brain_alterado.pd	109
Figura 5.29 – <i>Patch</i> write_pd.pd responsável pelo envio do valor do Pacote seleccionado no porto 8004	111
Figura 5.30 – Histórico do sistema	115
Figura 5.31 – <i>Wearable</i> afectivo	116

# 1 Introdução

Estudos científicos recentes indicam que as emoções desempenham um papel fundamental em todas as nossas tomadas de decisão, na percepção, na aprendizagem e num variado número de outras funções cognitivas. Desempenham um papel fundamental na motivação, influenciando fortemente a percepção, cognição e a criatividade (Picard, 1997a).

As emoções não se limitam às artes, ao entretenimento ou às interacções sociais; as emoções influenciam o próprio mecanismo do pensamento racional do indivíduo.

Sabe-se agora pela experiência que deverá existir um equilíbrio entre a emoção e o racionalismo: demasiada emoção ou emoções fortes podem afectar o mecanismo do pensamento racional, mas demasiado pouca emoção afecta também esses mesmos mecanismos racionais, impossibilitando um indivíduo de ter um pensamento lógico e inteligente. Torna-se assim necessário um equilíbrio saudável entre emoção e racionalismo de modo a que se possam tomar as melhores decisões, de modo a que se possam resolver os problemas de uma forma criativa e flexível.

Estudos científicos concluem que ter emoções não se trata de um “luxo”. A emoção tem um papel fundamental no comportamento racional e inteligente. Não só contribuem para uma interacção mais rica, mas influenciam também a forma como um indivíduo interage de uma forma inteligente

Mas e em relação aos computadores (quando se fala em computadores está-se a pensar em programas informáticos, agentes, etc.)? Não deveria a emoção ser completamente posta de parte, uma vez que esperamos desses mesmos computadores fiabilidade e um comportamento previsível?

Tal como a emoção é necessária no pensamento humano para a tomada de decisões, torna-se também necessária nos computadores se se quiser que sejam capazes de tomar decisões racionais. Desta forma, deverão possuir emoções (ou mecanismos semelhantes às emoções) associadas aos mecanismos puramente lógicos actuais.

Torna-se assim necessário, para que os sistemas informáticos em geral possam ser considerados “inteligentes”, para que se adaptem

aos utilizadores, para que naturalmente possam interagir, estarem dotados da capacidade de reconhecerem e expressarem emoções, possuírem emoções, e possuírem o que normalmente se denomina de “inteligência emocional”.

Neste âmbito, surgiu assim o conceito de Computação Afectiva: computação que se relaciona com, emerge de, ou deliberadamente influencia emoções (Picard, 1997a). A computação afectiva inclui a implementação de emoções, podendo ser uma ajuda importante no desenvolvimento e no teste de novas e velhas teorias das emoções. A computação afectiva inclui também muitos outros aspectos, tais como dotar um sistema informático com a capacidade de reconhecer e expressar emoções, desenvolver a sua capacidade de resposta inteligente a emoções humanas, e dar a capacidade a esses sistemas de regular e utilizar as suas emoções.

Como pode um computador reconhecer o estado emocional de uma pessoa? Como pode um computador utilizar os sinais bio-sintéticos de forma a parecerem apropriadamente emocionais? Estas são questões que cada vez mais investigadores têm colocado e para as quais ainda não existem respostas definitivas.

A Computação Afectiva, embora ainda seja uma área de pesquisa recente, pode ter um papel fundamental para o avanço na forma como compreendemos a emoção e a cognição, em melhorias na forma como os computadores resolvem os problemas, em avanços na forma como poderemos comunicar com esses mesmos computadores e na forma como influenciarão o próprio desenvolvimento humano.

O foco deste trabalho concentra-se assim no desenvolvimento de um sistema multimédia capaz de captar os sinais fisiológicos de um indivíduo e utilizar esses sinais para interagir em tempo real com conteúdos multimédia (imagem, vídeo, som e texto). Para isso, e com base nesses mesmos sinais fisiológicos, pretende-se reconhecer o estado emotivo do indivíduo, reflectindo-se esse estado também na interacção com todo o conteúdo do sistema.

Pretende espelhar emoções interiores, transmutando-as, provocar uma polifonia contínua de emoções que geram sentimentos e, por sua vez, novas emoções, transpondo-as para uma nova interpretação da consciência humana, e, neste fluxo contínuo, provocar o conhecer e o ser, criar espaço de reflexão aprofundada.

Os principais objectivos propostos para este trabalho foram:

- efectuar um estudo da área ainda recente da Computação Afectiva;
- estudar as principais ferramentas necessárias à recolha e processamento de sinais fisiológicos;
- definir os requisitos e a arquitectura do sistema capaz de utilizar os sinais;
- implementar um protótipo do sistema;
- aplicar o protótipo num *wearable* desenvolvido para o efeito;
- efectuar uma avaliação do sistema desenvolvido.

A estrutura da dissertação reflecte o percurso efectuado em termos do estudo considerado relevante para o projecto prático — a elaboração do protótipo —, isto é, que o fundamenta e o sustenta nas opções tomadas para a sua implementação.

Segue-se à introdução apresentada no presente capítulo, apresenta-se no segundo capítulo – Emoções e Computação Afectiva – um estudo sobre esta recente área de investigação. É também efectuado um estudo sobre os sinais fisiológicos de um indivíduo que servem de base à Computação Afectiva para a detecção de emoções.

No terceiro capítulo – Ferramentas – são analisadas as principais ferramentas em termos de *hardware* e *software* para a captação e o processamento dos sinais fisiológicos, assim como ferramentas e linguagens de programação multimédia para a implementação do protótipo que irá utilizar esses mesmos sinais.

No quarto capítulo – Definição do sistema e especificação – é feita a especificação do sistema e a definição da sua arquitectura em termos de interligação dos vários módulos de software e do fluxo de dados do protótipo.

O quinto capítulo – Detalhes de implementação e resultados – apresenta-se detalhadamente todas as fases da elaboração do protótipo, desde a concepção inicial até à integração. É também apresentado o *wearable* onde o protótipo foi aplicado. A finalizar este capítulo são apresentados os resultados de um inquérito realizado para obtenção de informação sobre o conteúdo multimédia do protótipo.

No sexto capítulo – Conclusão – são apresentadas as conclusões decorrentes deste trabalho de investigação e são lançadas as perspectivas de futuros desenvolvimentos, incidindo essencialmente no aperfeiçoamento do protótipo implementado.

## 2 Emoções e Computação Afectiva

O neurocientista António Damásio (Damásio, 2003) define «uma emoção [como sendo] uma colecção de respostas químicas e neuronais (...) automáticas, (...) [a resolução] de uma alteração temporária e profunda do estado do corpo e das estruturas cerebrais que mapeiam o corpo e suportam o pensamento. (...). [Na] fase de avaliação (...) [vislumbramos] a beleza e a espantosa inteligência representada pelas emoções, bem como a forma poderosa das emoções resolverem tantos dos nossos problemas.»

Se as emoções são os aspectos da realidade humana que nos conferem maior capacidade de percepção, reacção e adaptação, será que poderemos levar estas características para os computadores e tornarmos a sua interacção connosco mais rica e adaptativa, numa base eminentemente mais humana? Será que os computadores nos poderão transcender? Em que aspectos? As emoções reveladas pelos computadores serão limitadas ou poderão desenvolver-se livremente, à medida que eles as “sintam”? Sentir-nos-emos invadidos na nossa privacidade? Teremos mais um amigo? Onde nos levará um salto qualitativo tão grande?

Por conseguinte, os computadores terão que ser capazes de reconhecer e avaliar as emoções, cruzar os seus valores com outras capacidades afectivas, expressar emoções, desenvolver emoções e a própria inteligência emocional. Será que a tecnologia de que dispomos no momento nos permite implementar esta abordagem comportamental? Quais as ferramentas de que dispomos actualmente?

### 2.1 Emoções: fisiologia e cogniscência

Associada também ao sentido da visão, a percepção é influenciada pela emoção como em muitos outros processos inerentes ao Homem. Rosalind Picard, uma engenheira ligada à investigação nesta área, actualmente directora no M.I.T. (*Massachusetts Institute of Technology*) do grupo de Investigação de Computação Afectiva, afirma que «as emoções têm um papel importante no processo da tomada de decisão, percepção, aprendizagem e numa grande variedade de funções cognitivas.(Picard, 1997a)»

As emoções têm um papel catalisador, actuando na cognição, na imitação, na atenção, no planeamento, no raciocínio, na aprendizagem, na memória, na criatividade, na inteligência, na resolução de problemas, na tomada de decisão e na percepção.

As emoções conferem competências emocionais ao ser humano, dotando-o de um comportamento mais rico, em termos perceptivos, adaptativos e de resposta. O estudo, a compreensão e a documentação desta riqueza comportamental impõem-se para munir os computadores de uma maior inteligência e de um comportamento mais humano. Os processos mais eficientes têm-se manifestado como sendo os biológicos e humanos: a tecnologia procura estudar a realidade e imitar essa riqueza e adaptabilidade, como forma de melhorar a interface homem-máquina, ou seja muni-la de inteligência de resposta e até reveladora de emoções.

Em síntese, admite-se a existência de uma “inteligência emocional”, que terá que ser implementada no computador, para o aproximar do comportamento humano, o tornar capaz de reconhecer e demonstrar emoções, estender as suas capacidades em termos de interacção homem-máquina, e para o aproximar dos comportamentos humanos, quer em termos de compreensão ou mesmo de reacção.

Após vários estudos, o neurologista Richard E. Cytowic concluiu que as emoções estão relacionadas com um aumento de actividade das zonas do cérebro designadas por sistema límbico e que compreendem, aproximadamente, o hipotálamo, o hipocampo no lobo temporal e a amígdala. Este sistema é o responsável pela atribuição de uma valência (*valence*) emocional qualitativa aos acontecimentos (se se sentem como positivos ou negativos) e a que damos importância (e, que portanto, capta a atenção) (Picard, 1997a).

Ainda segundo este investigador, o sistema límbico contribui para a flexibilidade, para a imprevisibilidade e para a criatividade do comportamento humano, possuindo também vastas interconecções com o córtex, sendo as funções do cérebro consideradas uma mistura dos dois sistema límbico e córtex, interagindo e funcionando como um todo. Para o professor de psicologia norte-americano Carroll Izard, a emoção é uma «motivação e uma força que guia na percepção e na atenção (Picard, 1997a)».

Neste estudo conclui-se que a percepção não é apenas uma função puramente cortical. Ao nível da tomada de decisão, especialmente a



racional, é considerada como sendo uma função puramente cognitiva e de alto nível dentro do cérebro (Picard, 1997a). Actualmente, aceita-se que o cérebro tem uma componente racional e emocional, mesmo considerando que possui funções específicas e adequadas a situações muito díspares, quase opostas. A actividade cortical e a límbica complementam-se e completam-se, sendo difícil definir se apenas uma destas partes está a funcionar, e qual delas.

A tomada de decisão está, de uma forma geral, associada à actividade cortical, embora se tenha que admitir que possui os seus alicerces na actividade límbica. Geralmente, o pensamento encontra-se ligado à racionalidade e à actividade cortical, considerada de mais alto nível. Por outro lado, o sentimento está ligado às emoções e à actividade límbica. Nos estudos efectuados verificou-se que a maior parte das pessoas funciona ao nível do pensamento, e apenas uma pequena parte do sentimento, sendo estes por vezes associados com atitudes e comportamentos masculinos e femininos, respectivamente (Picard, 1997a).

Estudos mais aprofundados, sobre a actividade cerebral neuro-anatómica, revelaram que a zona do hipocampo é o local onde tudo converge. Todas as entradas sensoriais, externas ou viscerais, têm que passar pela parte límbica do cérebro antes de serem redistribuídos pelo córtex, para serem analisados, passando novamente para o sistema límbico, para sofrer uma avaliação, se a entrada multi-sensorial é importante (Picard, 1997a). Na realidade, o que se verifica é a necessidade de um equilíbrio entre o nível do comportamento emocional, sem reduzir a emoção nem a exacerbar.

Os computadores que implementam máquinas de inteligência artificial apresentam um grande predomínio de regras e factos, mas são relativamente ineficientes na tomada de decisão, uma vez que associam julgamentos de valores a acontecimentos que se notabilizam na tomada de decisão. Os julgamentos de valores decorrem da actividade e da interacção entre o sistema límbico e o córtex cerebral. Na cognição humana estão subjacentes aos processos da emoção, da percepção e da tomada de decisão. Para dotar os computadores destas apetências e torná-los mais eficazes, como sistemas cognitivos, é necessário estudar aprofundadamente estes fenómenos humanos, para os dotar de capacidade de decidir e de mecanismos típicos das emoções.

Com o objectivo de dotar o computador de emoções, mesmo através da comunicação textual, é necessário que o computador perceba e expresse emoções e que seja capaz de, pelo menos, utilizar o humor para expressar emoções, como sugeria Hofstadter (Picard, 1997a). O teste de Turing, um teste clássico para a avaliação da inteligência dos computadores, inclui som e fúria através de discurso verbal, utilizando o computador para expressar emoções (Picard, 1997a). Outro aspecto a estudar é a inteligência, que actualmente possui múltiplas formas, incluindo a inteligência social, sendo constituída por capacidades inter e intra-pessoais e, também, por inteligência emocional, definida por Salorey e Mayer como «a capacidade de monitorizar os sentimentos e emoções, de uns e de outros, e de os distinguir entre eles para usar esta informação e guiar o pensamento e as acções individuais (Picard, 1997a)».

### **2.1.1 Comunicação afectiva**

A inteligência emocional compreende aspectos como empatia, controlo dos impulsos, persistência e preocupação consigo próprio. O afecto é uma parte da comunicação, social e humana e pode por acréscimo ser utilizada na interacção de humanos com computadores (Picard, 1997a), sendo muitas vezes utilizado na literatura para conferir características humanas a objectos inanimados e animais.

A linguagem corporal é também uma forma de comunicar, de participar e reforçar a comunicação, conferindo-lhe, desta forma, mais elementos. Estes aspectos da comunicação e interacção humana devem ser levados para os computadores para lhes permitir reconhecer e expressar afectos.

Existem outros elementos corporais, como o batimento cardíaco, a respiração, a pressão sanguínea, a rigidez muscular e a postura, que dão informação acerca do estado emocional da pessoa e que permitem distinguir diferentes estados emocionais (Picard, 1997a).

Do estudo de investigação efectuado verifica-se que a emoção tem duas componentes fundamentais: a cognitiva, que reforça a sua componente mental; e a física, que reforça a sua componente corporal. A compreensão destes dois aspectos é importante para dotar os computadores de capacidade de reconhecer ou compreender os sinais externos da emoção, permitindo-lhes

identificar a emoção vivida pela pessoa através dos sinais exteriores, como por exemplo, a voz, a face, os gestos e avaliar a situação em que a pessoa se encontra.

Em 1977 surge pela primeira vez o termo *sentic*, criado pelo neurocientista e músico Manfred Clynes. O termo *sentic*, originado a partir da palavra em Latim *sentire*, tem a sua etimologia nas palavras sentimento e sensação, que trocam de significado com os adjectivos emocional e afectivo, reforçando os aspectos físicos da expressão da emoção. O estado emocional, segundo Rosalind Picard, refere-se às nossas dinâmicas interiores quando afectados pela vivência de uma emoção, tendo componentes do estado físico e emocional, e evolui no tempo com uma variedade de outros factores de activação e de condicionamento.

Uma experiência emocional refere-se à pessoa como um todo que se apercebe do seu próprio estado emocional. O termo expressão emocional será utilizado frequentemente para descrever o que é revelado aos outros, voluntária ou involuntariamente. A expressão emocional, através do sistema motor ou de outros sistemas corporais ou da modelação *sentic*, é geralmente involuntária e dá indicações às outras pessoas acerca do estado emocional que uma determinada pessoa está a viver (Picard, 1997a).

O estado de espírito (*mood*) também é despoletado quando uma emoção é repetidamente accionada estabelecendo o estado de espírito correspondente, sendo o que Manfred Clynes designava como modulação *sentic*, e expressa a componente física das emoções já comumente aceite pelos estudiosos da emoção e da neurobiologia. Assim, a modulação *sentic* é referida como uma inflexão de voz, uma expressão facial, uma postura ou seja, traduz-se no significado físico através do qual um estado emocional tipicamente se exprime, sendo este o primeiro significado da comunicação da emoção humana (Picard, 1997a). Este processo de comunicação é natural e funciona ao nível do subconsciente, portanto, de uma forma inata.

A fisionomia e o estudo dos aspectos faciais, para o Prof. Paul Ekman, da Universidade da Califórnia, em S. Francisco, indicam e traduzem as emoções básicas, e, embora cada pessoa tenha o seu próprio padrão de utilização dos músculos faciais, ele afirma que é provável que exista um padrão comum o que poderá simplificar a aprendizagem e o reconhecimento por parte do computador através

da modulação *sentic* operada no rosto do utilizador. Paul Ekman designou “regras de visualização social” e, mais tarde, criaria o “sistema de codificação das acções faciais”, que iria servir de suporte aos primeiros sistemas de reconhecimento automático de expressões, através do mapeamento realizado entre os músculos faciais e as emoções.

A entoação e o tom de voz fazem parte da modulação *sentic* e, portanto, traduzem e transcendem o conteúdo de uma mensagem verbal adicionando-lhe significado emocional. Nesta área o desafio para os computadores aumenta, nomeadamente no que concerne identificar o que é dito, quem emite a mensagem e como é dito.

Manfred Clynes explorou formas motoras de expressão, durante expressões emocionais que designou como de equivalência *sentic*, traduzidas em pressões e deflexões numa superfície flexível, sendo estes sinais, diferentes formas de comunicar emoções, por sua vez designados como formas *essentic*.

Existem várias respostas fisiológicas que combinadas podem ajudar numa avaliação mais correcta do estado emocional, entre elas o batimento cardíaco, a pulsação, a dilatação das pupilas, a respiração, a condutividade eléctrica, a cor e a temperatura da pele (Picard, 1997a).

Contudo, existem emoções que não terão formas *essentic* bem delineadas, definidas e de valores expressivos, como a vergonha e a culpa, comparativamente com as emoções de ódio e de alegria, mas que também têm as suas manifestações próprias, que se traduzem ao nível da postura corporal, gestos e comportamentos e, ainda, por alguns sinais involuntários como a dilatação das pupilas.

Está comprovado cientificamente pela neurobiologia, pela neuroquímica e pela psiconeuroimunologia que o estado emocional altera o aspecto fisiológico da pessoa, afecta o corpo como um todo, e a sua resposta à doença, nomeadamente ao nível dos linfócitos (capazes de matar os agressores) e dos macrófagos (capazes de absorver os intrusivos), impedindo o sistema imunitário de reagir e lutar pela sobrevivência contra as intrusões, nos processos de recuperação da saúde. Verificou-se também que as hormonas produzidas durante períodos de stress têm impacto nas células imunitárias (Picard, 1997a).

Manfred Clynes, depois do estudo das emoções e das repercussões no homem, chegou à conclusão de que existem vários factores de variabilidade do registo fisiológico expressivo da emoção no ser humano, nomeadamente no que diz respeito à intensidade da emoção vivida, ao tipo de emoção sentida, à forma como o estado emocional é induzido e às regras sociais de visualização, que instigam a expressão ou supressão de uma emoção sentida.

Mais problemas foram identificados no reconhecimento de emoções, tais como, alguns sinais fisiológicos de emoções que podem surgir não apenas pelas emoções vividas, mas por outras condições, produzindo um resultado fisiológico similar. Medicamentos, dietas, sono e hormonas podem alterar o estado de espírito da pessoa e influenciar a emoção que possa ser vivida ou induzida. O estado de espírito tem também influência sobre a cognição, traduzindo-se na «emoção a influenciar a cognição e a cognição a influenciar as emoções (Picard, 1997a).»

No decurso do levantamento de particularidades e obstáculos ao reconhecimento automático das emoções, surge o problema de construir modelos independentes da pessoa, que possam definir a forma de visualização de uma emoção tendo em consideração que a expressividade das emoções pode ser afectada por factores como temperamento, personalidade, género, profissão, idade, contexto social e cultural, entre outros.

Uma forma de contornar o problema da variabilidade da interacção humana será proporcionar uma fase de aprendizagem e reconhecimento entre o computador e o utilizador que, tal como no relacionamento entre humanos, pode ser facultada e propiciada ou poderá ser rejeitada, estendendo-se ao contexto onde o computador se encontra: este tipo de aprendizagem é designada por reconhecimento afectivo e significa tornar o computador perceptivo. Por outro lado, permite ao utilizador seleccionar o computador com o qual pretenda ter uma interacção mais profunda ao nível de revelar as suas emoções e de dispendir tempo a efectuar essa aprendizagem.

Construir um computador afectivo passará por equipá-lo com câmaras, microfones, sensores fisiológicos, ferramentas sofisticadas de reconhecimento de modelos, para efectuar o reconhecimento de sinais fisiológicos constituintes de uma emoção e inferir o estado emocional mais provável. Contudo, não é possível ainda determinar

o impacto das emoções noutro sistema corporal, tal como o sistema imunitário, embora estes sistemas de reconhecimento de emoções já se encontrem ao serviço da medicina preventiva (Picard, 1997a).

### **2.1.2 Aspectos cognitivos das emoções**

As emoções podem ser classificadas como primárias e secundárias. As primeiras designam as emoções sentidas de uma forma inata e emocional, accionadas directamente pelo sistema límbico, como o medo, por exemplo. As secundárias têm a ver com um estado cognitivo, que envolve reconhecimento, comparação, categorização, inferência, atribuição e julgamento. Surgem em estádios superiores do desenvolvimento humano, estabelecendo-se entre as emoções inatas e categorias de situações e objectos que as originam ou mesmo pensamentos cognitivos, actuando o sistema cerebral límbico-cortical e as suas ligações (Picard, 1997a).

Este último tipo de emoções corresponde a uma aprendizagem/cogniscência através de uma maturação típica da aquisição de consciência própria e só possível a partir da fase de resolução identitária, típica da adolescência, permitindo sentir emoções de culpa e de vergonha. Segundo Rosalind Picard, traduz-se em ligações límbico-corticais que funcionam em ambas as direcções «ligando pensamentos a sentimentos e sentimentos a pensamentos (Picard, 1997a)».

Os problemas que se colocam para a leitura e tradução fisiológica dos sinais de uma emoção, em particular num ser humano, também se colocam nos aspectos cognitivos de uma emoção. O contexto e a educação modelam e alteram o comportamento. A cogniscência do ambiente, mesmo quando controlada, pode ser difícil, pois implica análise visual e reconhecimento dos elementos do espaço, mantendo-se o problema de perder a riqueza do olhar humano que capta aspectos subliminares e de julgamento afectivo, filtrados e combinados com a sua história de vida, valores, aprendizagens e da própria maturidade emocional, mas que são em geral importantes para a cognição e compreensão da situação.

Investigando a história, algumas figuras proeminentes em termos de produção criativa, na literatura, na pintura, na psicologia, na composição musical ou na poesia revelaram um comportamento muito emotivo, chegando a depressões e até a actos de suicídio,

numa razão superior ao que se verifica na população comum. Do estudo de investigação emerge que a criatividade está associada a uma elevada emotividade, favorecida por um estado de espírito positivo, que facilita o reconhecimento de relações entre aspectos diferentes dos problemas, a descoberta de princípios para integrar e lembrar informação, a identificação de aspectos e características não usuais e a criação de associações e ligações completamente novas. Sempre que elevados níveis de emotividade se registam verifica-se uma predominância do sistema límbico em relação ao córtex cerebral, capaz de maior racionalidade e maior regulação.

A emotividade afecta a memória e, como é reconhecido, afecta também o córtex e o sistema límbico na forma como interagem e influenciam, apesar de ser conhecido que o sistema límbico possui um maior número de ligações a enviar informações para o córtex do que ao contrário. Acredita-se que a emoção é capaz de despoletar o ressurgimento de um evento passado há muito tempo e guardado na memória.

Estudos neurológicos concluíram que uma emoção verdadeira percorre um caminho no sistema motor diferente de uma emoção simulada, sendo designado o verdadeiro caminho, e os resultados fisiológicos também são diferentes, podendo ser reconhecidos pelas pessoas mais experientes.

As emoções também podem ser induzidas aproveitando o reforço que se verifica entre o corpo e a mente: os actores exploram muito bem estes aspectos para reforçar a sua personagem e induzir as emoções adequadas. A indução também pode ser subliminar ou pode ser apenas uma influência.

Depois de uma abordagem sobre a génese das emoções e das suas materializações fisiológicas no Homem não restam dúvidas de que, apesar de uma certa conotação negativa, associada aos comportamentos predominantemente emotivos, as emoções concorrem para a manutenção da vida, para a inteligência humana, para a interacção social, para a tomada de decisão racional, para a percepção e no limite para a sobrevivência do Homem.

Não existe dúvida de que será extremamente útil munir os computadores de capacidades afectivas, para melhorar o relacionamento e a compreensão entre Homem e a máquina. Será que poderemos efectivamente traduzir os processos emocionais

para a máquina? Deixaremos de sentir menos ao vermo-nos mapeados e reconhecidos no que de mais rico temos?

## 2.2 Computadores mais humanos

Quando se fala em computador, por vezes, poderemos não estar a referir-mo-nos a um computador no sentido tradicional do termo, podemos estar a designar sistemas que integram componentes com vários CPUs, *transputers*, agentes computacionais, robôs, *wearables* ou dispositivos computacionais, entre outros.

Por outro lado, as emoções e o estado emocional desenvolvem-se e maturam no ser humano. As crianças inicialmente manifestam expressões emocionais que denotam estados subjacentes, como por exemplo chorar, rir, berrar ou sorrir. Durante o seu desenvolvimento físico e emocional, aprendem a reconhecer as emoções nos outros e em si próprias, a reconhecer situações que podem despoletar emoções e a gerir as emoções, mesmo até a dissimulá-las ou a omiti-las de todo, quando atingem o estado adulto. Poderemos admitir que um computador poderá passar por um processo similar, começando por possuir um conjunto inato de capacidades/aptidões afectivas e ferramentas para em contínuo adquirir novas capacidades afectivas (Picard, 1997a).

Estes processos serão a base para dotar um computador da inteligência emocional tão necessária para interagir com humanos, sendo capaz de reconhecer, expressar e ter emoções e até de as regular. Existem várias abordagens para permitir que um computador tenha capacidades afectivas; uma, que parece razoável, é a de que os computadores devem ter emoções à medida que elas lhe são necessárias num determinado contexto.

Não poderemos esquecer que os computadores devem permanecer subservientes ao nosso controlo; no entanto, dotá-los de capacidades afectivas repercutir-se-á num desempenho muito maior em termos de tomada de decisão mais eficiente e flexível, ao adquirirem a capacidade para endereçar múltiplos aspectos de uma maneira inteligente e reflexiva, ao tornarem-se capazes de avaliar e determinar a importância e ou a pertinência de um acontecimento ou facto, ao passarem a estar dotados de percepção e atenção mais humanas e muitos outros aspectos das interacções ao nível de processos de regulação ou de cognição (Picard, 1997a).



### 2.2.1 Computadores que reconhecem emoções

Para que os computadores adquiram as competências de reconhecer emoções será necessário apetrechá-los de equipamentos extra que lhes traduzam o nosso mundo sensório-motor, por exemplo, terão de ter dispositivos de visão e audição e se possível outros, tais como de leitura por infravermelhos da temperatura, medição da respostas electrodérmica, medição do ritmo cardíaco, entre outros, que lhes permitem reconhecer as emoções e inferir o estado emocional. Mediante os valores aferidos poderão reagir adequadamente, procurando uma resposta adaptativa e inteligente na interacção com a pessoa.

Só com o sentido da visão e da audição será possível reconhecer expressões faciais, gestos, linguagem corporal, tons vocais e inflexões na voz, determinando a emoção expressa. São designadas emoções públicas aferidas a partir de formas de modulação *sentic*. Esta faculdade nos seres humanos designa-se inteligência emocional. Neste tipo de reconhecimento existem alguns factores que devem ser tidos em consideração, como, por exemplo, a capacidade de aprendizagem de reconhecimento emocional, aspectos dependentes da pessoa em observação, capacidade de esconder o estado emocional real, entre outros. Estas capacidades nos computadores podem ser estendidas a outros níveis não possíveis na vertente humana, como, por exemplo, a visão infravermelha, capacidades auditivas fora dos valores normais para os humanos, ondas cerebrais, electromiograma, resposta electrodérmica, feromonas e pressão sanguínea. A leitura e o cruzamento destes valores podem levar ao reconhecimento de estados emocionais que nem sequer têm designação actualmente (Picard, 1997a).

É indubitável a capacidade computacional para gerir grandes quantidades de informação em tempo útil, bem como para recolher dados de que um ser humano sem qualquer apetrechamento computacional (um *wearable*, por exemplo) não conseguiria ter leitura, sequer. Aos computadores é possível usar perspectivas independentes ou dependentes da pessoa, realizar aprendizagens afectivas com autonomia (realizando essa aprendizagem sempre que necessite) ou sem autonomia, utilizando ou não o factor estado de espírito para o processo de avaliação ou reconhecimento.

Para o reconhecimento computacional das emoções existem vários aspectos tipicamente humanos a ter consideração e, outros tipicamente das máquinas, que permitem estender e superar as capacidades humanas a níveis não imaginados. Na Figura 2.1 estão ilustrados os critérios utilizados nesse reconhecimento (Picard, 1997a).

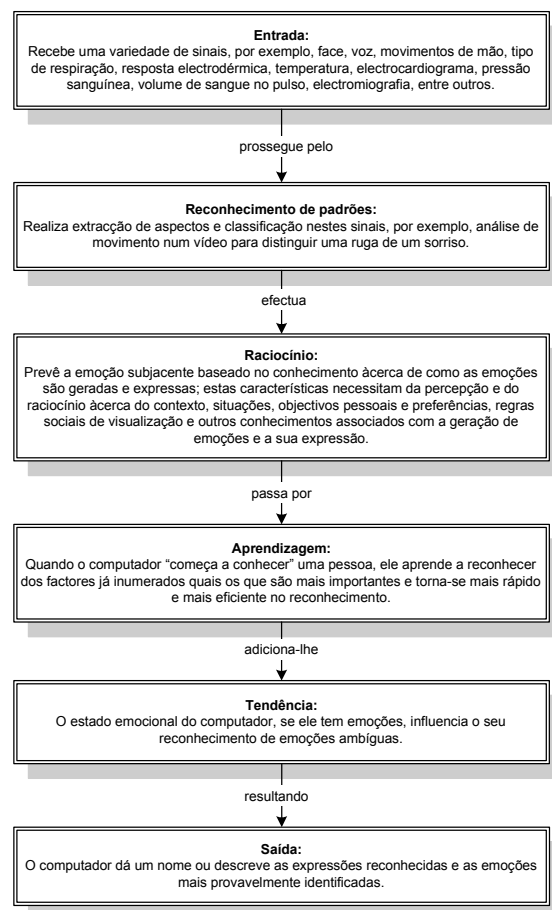


Figura 2.1 – Critérios utilizados no reconhecimento emocional por parte do computador

## 2.2.2 Computadores que exprimem emoções

Na comunicação verbal é frequente que o estado emocional seja revelado sem uma atitude racional ou mesmo até sem uma atitude consciente na recepção, mas adiciona-se ao conteúdo da informação transmitida enriquecendo-a. Esta informação adicional tem tradução na entoação, na altura e inflexões da voz. Por sua vez, a forma como o computador se expressa na sua interacção com o homem é uma forma que pode tornar a comunicação mais agradável e mais eficaz.

O modelo de Shannon e Weaver da comunicação foi estendido e, agora, falamos de “capacidade afectiva do canal” e largura de banda afectiva nas diferentes formas de comunicação, para reflectir a quantidade de informação afectiva que é transportada num canal, durante uma comunicação (Picard, 1997a). É reconhecido que consoante o tipo de canal de comunicação utilizado, pode passar mais ou menos informação afectiva, isto é, este pode possuir maior ou menor largura de banda afectiva.

Do estudo de investigação efectuado, conclui-se que a tecnologia e os ambientes mediados por computador dotam os canais de comunicação, e a própria comunicação, de maior largura de banda afectiva para expressar afecto e de outros canais diferentes, dos convencionais, existentes na comunicação inter-pessoal. As novas potencialidades emergentes do uso da tecnologia terão que ser criteriosamente utilizadas pois nem sempre serão úteis numa comunicação as novas capacidades disponíveis (Picard, 1997a).

Outro aspecto explorado pela capacidade de um computador poder manifestar emoções é a possibilidade de influenciar o estado emocional dos seus utilizadores através da sua inteligência emocional e do seu estado emocional no momento, negativa e positivamente, como um certa forma de contágio ou de sinergia indutora. Muitos aspectos têm que ser estudados, tais como os computadores podem encontrar-se num estado emocional interno e expressar um estado emocional completamente diferente e, por exemplo, qual a dose de exagero que se deve inferir de um determinado estado emocional para influenciar emocionalmente o interlocutor e tornar clara a emoção que quer expressar.

Utilizam-se alguns critérios para permitir aos computadores expressar emoções estabelecendo interfaces empáticas e relacionamentos de grande inteligência emocional, ilustrando-se na Figura 2.2 os critérios para a expressividade (Picard, 1997a).

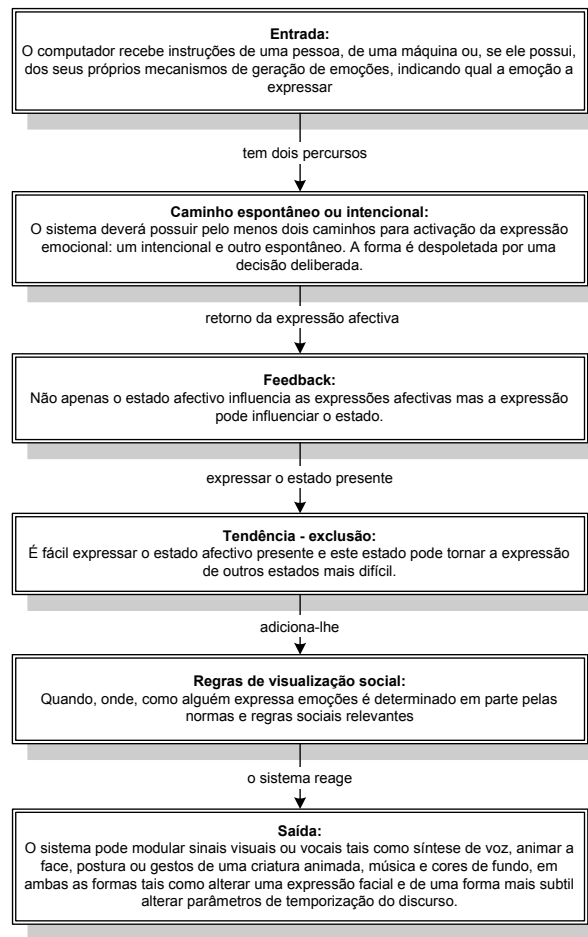


Figura 2.2 – Critérios de expressividade de emoções

### 2.2.3 Computadores que possuem emoções

A discussão sobre a temática se os computadores podem ter emoções e mesmo sentimentos não está concluída. Rosalind Picard tem-se dedicado muito a esta temática e sugere uma aproximação simplista como abordagem de classificação. Para esta investigadora, um sistema que comporte cinco componentes que estabeleceu, baseando-se nos sistemas humanos e biológicos, é considerado um sistema que “tem emoções”, embora possam não estar todos activos em simultâneo mas apenas alguns (Picard, 1997a).

Este sistema é constituído pelos seguintes componentes: emoções emergentes e comportamentos emocionais, emoções primárias rápidas, emoções geradas cognitivamente, experiência emocional e interações corpo-mente. Estes componentes foram extrapolados de um sistema emocional humano completo.

A componente das emoções emergentes é a primeira componente a ser atribuída a sistemas baseados nos comportamentos emocionais observáveis, nomeadamente quando o sistema que está a revelar o comportamento não tem mecanismos internos explícitos ou representações das emoções. Nos humanos, este sistema refere-se a comportamentos emocionais e expressões típicas da sua existência. A capacidade de expressar emoções pode ser vista como um tipo de comportamento emocional e até como podendo ter emoções atribuídas a ele.

A segunda componente, designada emoções primárias rápidas, encontra-se associada a reacções inatas e rápidas que estão ligadas a emoções de medo, surpresa e raiva, segundo o neurocientista António Damásio. O estudo de investigação do circuito destas emoções revela que as emoções primárias seguem dois caminhos nos sistemas de comunicação – um sistema de reconhecimento de padrões rudimentar (que actua rapidamente e pode bloquear o córtex, embora muitas vezes se engane, despoletando falsos alarmes) e um sistema de reconhecimento de padrões (mais apurado, mas mais lento e mais preciso). Funcionando em unísono, detectam eventos importantes, accionando os sentimentos subjectivos, a modulação *sentic*, os controlos de regulação e os comportamentos importantes no sentido da sobrevivência.

As emoções geradas cognitivamente constituem a terceira componente do sistema de reconhecimento emocional, para reforçar a explicação da geração das emoções. Acredita-se que estas emoções são originadas no córtex, que surgem lentamente e podem ser despoletadas apenas por pensamentos. António Damásio classificou estas emoções como “emoções secundárias”, caracterizadas por serem originadas no córtex e subsequentemente activarem respostas límbicas e sentimentos corporais, traduzidos no que se designa como numa experiência emocional com sentimentos subjectivos, principalmente se a emoção for intensa (Picard, 1997a). Verificou-se que para sistemas de inteligência artificial (AI) é importante que os computadores tenham capacidades de justificação das emoções, em sistemas que interagem com humanos, como por exemplo, resolução de problemas cooperativos e para a compreensão da linguagem natural. Num ser humano a existência de apenas emoções de origem cognitiva traduzir-se-ia numa deficiência; nos computadores é o processo mais frequente de lhes dar origem (Picard, 1997a).

O quarto componente traduz-se na experiência emocional, com tradução em três componentes fundamentais: a consciência cognitiva do estado emocional, a consciência das alterações fisiológicas que acompanham o estado emocional e o sentimento subjectivo interno que assiste na tomada de decisão e permite ter opinião acerca de acontecimentos, de agrado ou desagrado. Um sistema terá que ser munido com a capacidade de ter consciência cognitiva do seu próprio estado emocional pois é determinante para a experiência emocional. A consciência das alterações fisiológicas sofridas pela vivência de um estado emocional também é importante, mesmo para os computadores, cuja fisiologia difere grandemente da fisiologia humana, permitindo dotá-los de mais informação interna.

O último aspecto da experiência emocional é o sentimento subjectivo interno que nos permite ter opiniões de agrado ou desagrado e que será difícil implementar numa máquina pois traduz factores subjectivos da experiência humana que nos apoiam continuamente nas decisões que tomamos (Picard, 1997a). Segundo Rosalind Picard,

«uma experiência emocional dá-nos a capacidade de compreender melhor e aprender como regular as nossas experiências – por exemplo, porque é que queremos determinadas coisas, o que correspondeu ou não às nossas necessidades e como poderemos fazer as coisas melhor. (...) Através da experiência emocional nós ganhamos visão dentro das nossas próprias motivações e valores; tornamo-nos capazes de compreender melhor e utilizar a poderosa influência que as emoções exercem (Picard, 1997a).»

A última componente reporta-se às interacções corpo-mente. Ao longo do estudo das emoções ficou claro que elas influenciam a tomada de decisão, a percepção, o interesse, a aprendizagem, as prioridades, a criatividade e muitos outros aspectos do comportamento humano, nomeadamente a cognição e a inteligência, bem como, outros sistemas fisiológicos para além do cérebro, tais como as expressões faciais, a entoação vocal, a postura, os movimentos, os processos digestivos e até o funcionamento do próprio sistema imunitário (Picard, 1997a). Ficou explícito que as

emoções monitorizam um comportamento intrincado e simbiótico entre o corpo e a mente.

Nos computadores que não possuem um corpo humano, o que poderemos fazer é imitar a influência da emoção, podendo afectar e modular os modos que ele possua para se exprimir, imitar a modulação *sentic*, mas não poderão imitar os efeitos eletroquímicos ou bioquímicos, nem mesmo o sistema imunitário; poderão sim desenvolver outras reacções através de equipamentos e programação adicional que dote a máquina de capacidades de acção e reacção em tempo real e segundo determinadas restrições, funcionando mais numa base de sistema de controlo análogo aos mecanismos de regulação accionados pelas emoções.

O que importa reter é a necessidade de construir num computador todas as funções e as influências de uma emoção, sem efectivamente construir a emoção em si mesma. Resta-nos a incerteza sobre se conseguiremos implementar num computador de uma forma flexível e eficaz os mecanismos e as influências das emoções não estaremos efectivamente a construir um sistema emocional? A dúvida e o receio de, se de alguma forma, os sistemas e os equipamentos se aperfeiçoarem bastante e o suficiente, sermos suplantados por máquinas que nós próprios construímos?

Como síntese apresenta-se um resumo dos principais componentes de um sistema emocional, ilustrado na Figura 2.3 (Picard, 1997a).

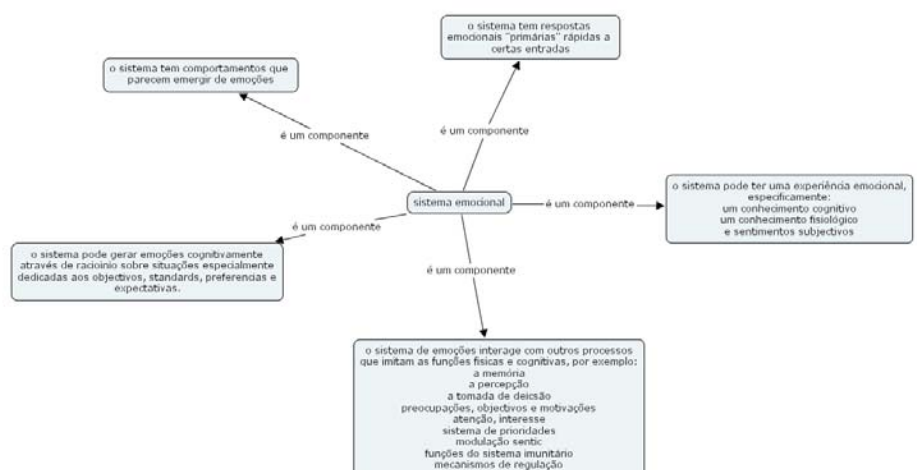


Figura 2.3 – Componentes de um sistema emocional

É tentador pensar que, como o sistema límbico é responsável por emoções mais imediatas e até passíveis de traduzir erros, o sistema

cortical é capaz de processar as emoções de uma forma mais lenta mas mais segura e adaptativa, sendo o software no computador o equivalente ao processamento cortical e o *hardware* o equivalente ao processamento límbico. Esta analogia simplista pode deixar-nos frustrados quando se verifica que a distinção entre *hardware* e software se dilui quando se tenta implementar um sistema emocional num computador ou agente computacional através de software (Picard, 1997a).

Só por software será possível implementar um sistema emocional, bem como um sistema de auto-sobrevivência que entre em funcionamento sempre que detecte que a sua existência está ameaçada, bloqueando todos os outros processos e desenvolvendo mecanismos e acções de recuperação das falhas detectadas. Estes tipos de sistemas são desenvolvidos e utilizados para recuperação a partir de falhas de *hardware*, de um disco ou *drive*, de danos físicos ou mesmo a partir da detecção da acção um vírus. São exemplos as máquinas Tandem, robôs e agentes de software.

A forma de um computador ter um corpo poderá ser traduzida num robô munido de sensores que implementem o órgão sensorial da pele, a dor, algoritmos de aprendizagem, mecanismos de leitura e tradução do que sente e do que o rodeia, permitindo influenciar o seu estado emocional e controlar as suas emoções aprendendo a evitar o que o possa destabilizar. A modulação *sentic* pode ser utilizada em algumas das suas formas humanas estendendo as possibilidades comportamentais de um computador.

O estudo que Rosalind Picard efectuou sobre as emoções e a consciência levaram-na a concluir que existem emoções, as primárias, que não estão subordinadas a nenhum acto de consciência e que existem emoções que implicam a existência de mecanismos de acção da consciência como é o caso das emoções de culpa e vergonha. Estas últimas emoções foram designadas como “emoções da auto-consciência”. É possível dar um pouco de consciência própria a um computador através de atribuição de um nome para onde se encontram, sobre o que estão a fazer, qual o estado em que se encontram afectivamente, o que é que os seus sensores estão a ler, entre outros valores que lhe podem ser dados, podendo mesmo deter um estado interno subjectivo, que comportará sempre diferenças significativas para o que possa ser um estado interno subjectivo humano (Picard, 1997a).



No que concerne ao estado emocional e à própria inteligência emocional, os computadores têm capacidade de expressar e reconhecer emoções, em si próprio e nos outros, regular as emoções próprias ou dos agentes com quem interagem e, por último, utilizar as emoções com o respectivo valor de motivação e de indução do estado de espírito. A capacidade de mostrar empatia também é importante na interacção com humanos. Estes processos passam por analisar interacções decorridas e melhorar futuros processos de interacção. Não poderão ser esquecidos os aspectos fundamentais de um computador, forte em memorização e armazenamento de informação, quer de factos, listas ou mesmo de situações imaginárias, mas ineficiente a adquirir valores e valências a partir das situações, e fraco em raciocínios ou analogias sempre que envolvam sentimentos, isto é, um significado emocional. Esta parece uma área ainda a necessitar de grandes desenvolvimentos (Picard, 1997a).

Quando se fala em interacção é importante distinguir quais os seus intervenientes uma vez que estes implicam que interacção possa ter características profundamente diferentes. Assim, a interacção computador-computador terá as suas próprias regras, não necessariamente inteligíveis pelo homem. A interacção computador-homem terá que integrar desde a inteligência emocional humana até os aspectos da máquina e implementar numa plataforma de “saúde emocional computacional”, que poderá ser gerida em contínuo como uma forma de aprendizagem, para aprender a ler o estado emocional do interlocutor e ser capaz de prever como poderá obter uma melhor resposta na interacção. Por último, a interacção humano-humano tem as suas próprias regras. Pode-se concluir que o sucesso da implementação das características afectivas tem a ver com o conhecimento dos requisitos para o domínio de aplicação específico.

Muito trabalho há ainda a fazer neste sentido, mas admite-se que se está a caminhar no sentido certo, permitindo que um sistema computacional que pode conter vários processadores e vários caminhos para comunicação, muitos mecanismos de regulação e múltiplos sensores, que pode ter mecanismos semelhantes aos das emoções nos computadores, consiga, mesmo através de um processo de duplicação, ter uma certa consciência das sensações e portanto consiga experiências emocionais. Neste momento, segundo Rosalind Picard, este último aspecto, a emoção, ainda é um mistério.

## 2.3 Aplicações da computação afectiva

Por se tratar de uma área de investigação recente, são muito poucos os exemplos práticos disponíveis da aplicação da computação afectiva. Embora seja ainda uma área dominada por projectos de investigação, adivinham-se já nos projectos actuais um variado número de aplicações que podem beneficiar da introdução de capacidades afectivas nos computadores. Uma das áreas de aplicação onde se tem verificado um maior investimento em termos de investigação e desenvolvimento de aplicações é na utilização de *wearables*.

### 2.3.1 *Wearables* afectivos

Uma das principais características que diferencia um sistema *wearable* de um simples computador ou equipamento informático portátil é a de que um sistema *wearable* pode estar em contacto com o corpo humano do utilizador durante um longo período de tempo sem que se torne incómodo ou condicionador da liberdade do utilizador. Por outras palavras, “um *wearable* é um sistema computacional que se pode vestir tal como uma peça de roupa ou de joalharia” (Picard, 1997b). Assim, um *wearable* pode ser um pequeno aparelho preso ao cinto de um utilizador, aplicado nos seus sapatos, luvas, chapéu, ou no vestuário, proporcionando variadas formas de contacto além do tradicional contacto da ponta dos dedos num teclado e no rato. Quando apetrechados de sensores especiais, capazes de captar os sinais fisiológicos do utilizador, e ferramentas destinadas ao processamento de sinal e reconhecimento de padrões evidenciados por esses sinais, um computador *wearable* pode, potencialmente, aprender a reconhecer e a detectar estados emotivos relacionados com os sinais fisiológicos de um indivíduo. A captação de sinais fisiológicos e o reconhecimento de padrões relacionados com esses mesmos sinais não é uma actividade recente; de facto, variados dispositivos electrónicos suficientemente pequenos para serem transportados pelo utilizador já se encontram disponíveis com o intuito de monitorizar alguns sinais vitais, tal como o batimento cardíaco, entre outros.

Um *wearable* afectivo possui um ponto em comum com os *wearable* destinados a fins medicinais ou de monitorização de sinais vitais: ambos permitem a captação de sinais fisiológicos do utilizador. Os

sistemas *wearable* destinados à monitorização de sinais fisiológicos são tipicamente utilizados para a monitorização de estados de stress e de outras manifestações corporais e emotivas, proporcionando a captação de dados à medida que o utilizador realiza as suas actividades diárias normalmente e sem interferências por parte do *wearable*. A captação de sinais fisiológicos para o reconhecimento de uma situação de stress possui importantes implicações para o desenvolvimento de interfaces homem-máquina inteligentes e funcionais, existindo actualmente uma tendência nas ciências dos computadores para o desenvolvimento de sistemas que aprendem aquilo de que utilizador necessita, tentando modelar os seus interesses. Contudo, uma forma natural do utilizador exprimir o seu agrado ou desagrado perante determinada situação e em contacto com um determinado sistema informático é através da expressão afectiva. Os utilizadores podem falar com uma voz calma ou irritada, sorrir ou mostrar uma cara de tristeza, podem gesticular de várias formas, sendo todas estas manifestações físicas uma indicação da forma como se sentem face a determinada situação ou em determinado momento (Picard, 1997b).

Um dos grandes problemas dos computadores actuais é o de serem “cegos” perante todas as “pistas” afectivas transmitidas inconscientemente pelo utilizador. Exemplos típicos de interacção através de menus apenas obtêm do utilizador respostas pré-definidas nesse menu, tal como se o utilizador se encontra satisfeito ou não com determinada resposta do sistema. Uma outra forma mais natural e fiável da detecção do sentimento do utilizador quando confrontado com determinada situação seria o da detecção do seu estado emotivo com base na sua expressão corporal. Um *wearable* afectivo possui assim uma oportunidade sem precedentes para o conhecimento do utilizador e de como este se sente através da captação da sua expressão corporal e da manifestação dos seus sinais fisiológicos.

Existe actualmente um número bastante elevado de aplicações da Computação Afectiva e dos *wearables* afectivos, para além das aplicações com fins medicinais já citadas (Picard, 1997b). Um exemplo de uma aplicação que se pode dar a um *wearable* afectivo é o de um leitor de CDs portátil e inteligente capaz de fazer a selecção das músicas com base no estado emotivo do utilizador. A música é considerada como a mais popular forma de manipulação do estado emotivo de um indivíduo. E embora seja praticamente impossível

adivinhar exactamente qual a música que determinada pessoa gostaria de ouvir em determinado momento, não é difícil decidir qual o género musical que poderá preferir em determinado momento consoante o seu estado emotivo. À medida que os *wearables* afectivos vão possuindo uma maior capacidade quer em espaço de armazenamento de músicas quer na capacidade para a detecção do estado emotivo e para que analise a resposta do utilizador (*feedback*), existe assim a oportunidade de aprendizagem e reconhecimento de padrões afectivos, recorrendo ao seu estado emotivo, a sinais fisiológicos, ao ambiente e às suas preferências musicais. Assim, este *wearable* afectivo pode reproduzir músicas autonomamente, tendo não só em conta as preferências musicais do utilizador mas também ter em conta o seu estado emotivo e o ambiente que o rodeia (Picard, 1997b).

Um outro exemplo de aplicação de um *wearable* afectivo é o do alargamento das fronteiras de expressão e comunicação humanas através de sistemas informáticos inteligentes e capazes de reconhecerem o estado emotivo do utilizador. Com base no seu estado emotivo podem desempenhar um papel de mediadores da interacção e comunicação humanas através da apresentação e manipulação interactiva de diversos media (imagem, vídeo, texto e som) que pretendem representar o estado emotivo do seu utilizador. Estes sistemas pretendem assim ser um espelho das emoções interiores, transmutando-as, provocando uma polifonia contínua de emoções que geram sentimentos e, por sua vez, novas emoções, transpondo-as para uma nova interpretação da consciência humana e, neste fluxo contínuo, provocar o conhecer e o ser, criar espaço de reflexão aprofundada (Picard, 1997b).

### **2.3.2 Outras aplicações**

Apresentam-se de seguida outros projectos relevantes relacionados com a aplicação da computação afectiva (MIT, 2004b):

- Brinquedos afectivos – o Tigre Afectivo é um brinquedo que reage de acordo com a emoção detectada no utilizador;
- Tutor afectivo – um agente em *software* orientado para o ensino que detecta o estado emotivo do aluno (ansiedade, tédio, interesse) e, de acordo com esse estado, é capaz de

ajustar o seu comportamento (como por exemplo, a velocidade de apresentação de informação);

- Resposta através do computador à frustração do utilizador – implementação de um agente para a interacção homem-máquina capaz de auxiliar o utilizador na recuperação de estados emotivos negativos, mais propriamente a frustração;
- Afectividade do condutor – projecto em que se pretende dotar um automóvel e um condutor de sensores para o reconhecimento de estados emotivos do condutor, tais como a raiva ou o stress;
- Reconhecimento da emoção num actor – reconhecimento de padrões correlacionados com os sinais fisiológicos do seu corpo durante a sua representação de estados emotivos;
- Óculos de expressão – um dispositivo *wearable* que permite a detecção do nível de interesse ou confusão por parte do utilizador através de técnicas de reconhecimento de padrões com base na captação dos movimentos dos músculos faciais;
- Espelho afectivo – um projecto que tem como objectivo a implementação de um agente em software capaz de detectar e responder em tempo-real ao estado emotivo do indivíduo, imitando a sua expressão facial e postura. Este projecto será integrado no Sistema de Ensino Afectivo;
- Sistema de Ensino Afectivo – projecto tendo em vista a criação de agentes afectivos que podem comportar-se como tutores inteligentes, companheiros virtuais, ou um grupo de amigos virtuais para apoio à aprendizagem e criatividade;
- BioMod – interface desenvolvida para utilizadores de *wearables* e dispositivos móveis capaz de detectar vários sinais fisiológicos com vista à apresentação de *feedback* sobre informação médica. Uma aplicação em desenvolvimento que envolve o BioMod é um sistema de monitorização do stress em indivíduos que estão a tentar deixar de fumar: quando se detecta um elevado nível de stress é enviado um alerta para um assistente médico para que lhe possa prestar apoio psicológico num momento de fraqueza;

- Cadeira de reconhecimento da postura – sistema que reconhece em tempo real padrões relacionados com a postura de um indivíduo através de sensores de pressão colocados numa cadeira. O sistema detecta estados emotivos nas crianças em situações de aprendizagem, detectando o seu nível de interesse, apatia ou irritação;
- Análise da interacção com um rato – este projecto tem como objectivo o desenvolvimento de ferramentas de detecção e adaptação ao estado emotivo do utilizador através da utilização de um rato. O rato modificado inclui um sensor de pressão para a detecção da valência emotiva do utilizador (gosta/atração ou não gosta/repulsa).

## 2.4 Sinais Fisiológicos

O objectivo da captação dos sinais fisiológicos de um indivíduo prende-se com a obtenção de uma melhor compreensão das relações e interacções entre esses mesmos sinais e o comportamento e estado emocional evidenciado pelo indivíduo na altura da sua captação. Estes sinais fisiológicos, estando relacionados com a actividade comportamental do organismo, são uma forma privilegiada de enriquecimento da informação relativa à relação entre a manifestação fisiológica de sinais observáveis e o estado emotivo do indivíduo, podendo-se assim estabelecer relações directas entre a sua actividade fisiológica e as suas emoções. Os sinais podem transformar-se progressivamente em elementos de caracterização, categorização e predição do estado emotivo do indivíduo.

Tendo por base o trabalho de Paul Ekman sobre a expressão facial, D. Christopher Dryer levou a cabo uma série de experiências sobre “em que medida os sinais fisiológicos poderiam ser utilizados para distinguir vários estados emotivos do ser humano” e concluiu que 4 sinais fisiológicos são suficientes para determinar eficazmente o estado emotivo de um indivíduo. Os sinais utilizados nas suas experiências foram a condutividade eléctrica da pele (GSR – *Galvanic Skin Response*), o batimento cardíaco (HR – *Heart Rate*), a temperatura da pele e a actividade somática geral (GSA – *General Somatic Activity*) (Dryer, 1999).

Além destes 4 sinais, existem muitos outros sinais fisiológicos relacionados com as respostas emotivas de um indivíduo que podem ser fisicamente adquiridos e medidos através de vários dispositivos ou sensores, tendo estes últimos a vantagem de poderem estar em contacto físico com o indivíduo através de uma forma não invasiva. Uma vez que o ser humano se encontra cada vez mais familiarizado com o contacto e o manuseamento de equipamentos informáticos, aumentando este contacto através de sensores possibilita-lhe uma nova forma de comunicação com esses mesmos equipamentos, sem um grande esforço de aceitação e adaptação por parte do utilizador.

Rosalind Picard tem vindo a realizar a sua investigação também com base em 4 sinais, mas diferentes dos sinais sugeridos por Dryer: GSR, pressão sanguínea (BVP – *Blood Volume Pressure*), actividade muscular (EMG - *Electromyogram*) e o sinal relativo à respiração (Picard, 1997a). Estes sinais podem muitas vezes, em ambientes laboratoriais e controlados, ser substituídos por apenas 3 quando não se possuem os recursos em termos computacionais ou de sensores para se fazer a captação simultânea dos 4 sinais referidos. Uma vez que o sinal relativo à respiração está intimamente ligado à frequência do batimento cardíaco, o sinal da respiração pode ser excluído sem perdas significativas em termos da exactidão dos resultados, sendo possível, como se demonstrará no protótipo desenvolvido no âmbito desta Dissertação, a detecção do estado emotivo de um indivíduo com base em apenas 3 sinais fisiológicos: HR, GSR e EMG.

As etapas para o reconhecimento das emoções com base em sinais fisiológicos são as que se apresentam na Figura 2.4.

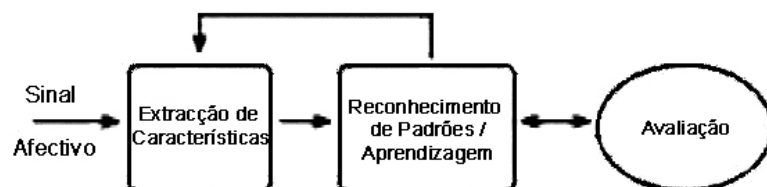


Figura 2.4 – Etapas do reconhecimento das emoções (Psychology, 2004)

A extracção de características é a área onde actualmente se encontra a maior parte da investigação relacionada com a computação afectiva, pretendendo-se determinar quais são as

características relevantes dos sinais afectivos para enviar ao módulo de aprendizagem. Uma vez que o módulo de aprendizagem actua com base nessas características e não nos sinais, é possível, por exemplo, utilizar módulos de aprendizagem desenvolvidos para actuarem em sinais visuais e aplicar as suas funcionalidades no processamento de sinais afectivos (Psychology, 2004).

### **2.4.1 Batimento cardíaco**

Ao registo da actividade eléctrica no coração dá-se o nome de Electrocardiograma (ECG). O princípio básico do registo do electrocardiograma é o da captação ao nível da superfície cutânea dos potenciais eléctricos gerados pela actividade do músculo cardíaco. Num coração saudável esta actividade de contracção e relaxamento do músculo cardíaco realiza-se de uma forma coordenada e cíclica, criando o que se chama de batimento ou ciclo cardíaco.

Para se entender o ciclo cardíaco e os mecanismos de sincronismo associados, é preciso, antes de mais, descrever os dois fenómenos que acontecem durante o batimento cardíaco. O primeiro é um potencial eléctrico que é gerado pelo nodo sino-atrial e o segundo é o potencial mecânico gerado pela contracção ordenada do miocárdio. No primeiro evento ocorre um disparo de actividade eléctrica que rapidamente se dissipa por cada célula miocárdica, o segundo evento é a resposta do conjunto de células miocárdicas que produz a contracção e por conseguinte o movimento de sangue dentro de cada câmara cardíaca.

O ciclo cardíaco é assim composto por dois eventos: a diástole e a sístole (ICB, 2004):

- a diástole é o enchimento das cavidades cardíacas com o volume de sangue;
- a sístole é a expulsão do sangue das cavidades cardíacas. Existem dois tipos de sístole, a atrial e a ventricular. Cada uma é precedida por uma diástole.

Na Figura 2.5 apresentam-se as fases do batimento cardíaco (ICB, 2004).



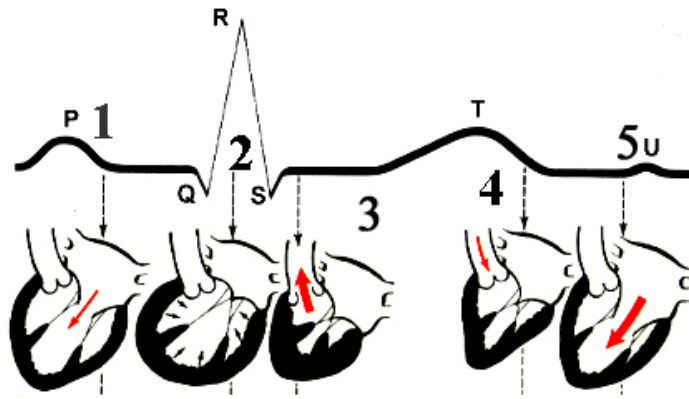


Figura 2.5 – Fases do batimento cardíaco (ICB, 2004)

1. Início da diástole, abertura das válvulas tricúspide e mitral e enchimento ventricular;
2. Fecho das válvulas de entrada, final da diástole;
3. Contração ventricular, abertura da válvula pulmonar e aórtica - sístole ventricular;
4. Final da sístole ventricular, fecho das válvulas pulmonar e aórtica;
5. Reinício da diástole atrial e ventricular

O electrocardiograma normal é composto por um conjunto de ondas tipicamente denominadas por P, Q, R, S e T. A onda P é produzida pelos potenciais eléctricos associados à contração auricular, enquanto que o complexo QRS está associado à contração ventricular e aos potenciais eléctricos por ela gerados. Por sua vez, a onda T corresponde à repolarização ventricular, representando o período de recuperação do músculo ventricular após a sua estimulação (Rodrigues, 1989).

A Figura 2.6 apresenta-se um exemplo das ondas P, Q, R, S e T do batimento cardíaco.

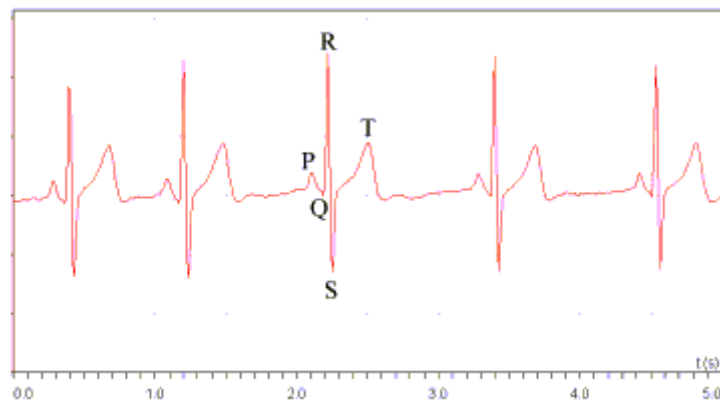


Figura 2.6 – Ondas P, Q, R, S e T do batimento cardíaco (Pursuit-Performance, 2004)

Assim, o ECG é composto por ondas de despolarização e repolarização. O fenómeno de despolarização deve sempre ocorrer para que ocorra o fenómeno de contracção. As ondas P ocorrem imediatamente antes do início de contracção atrial e o complexo QRS imediatamente antes da contracção ventricular. A repolarização atrial ocorre quase que simultaneamente com a despolarização ventricular e, por isso, não é observada no ECG. O ciclo cardíaco é definido como o intervalo entre as ondas R (Rodrigues, 1989).

Na Figura 2.7 apresenta-se o registo típico de um ECG.



Figura 2.7 – Exemplo de um registo de ECG (Seps, 2001)

## 2.4.2 Condutividade eléctrica da pele

A condutividade eléctrica da pele derivada da sua sudação (GSR – *Galvanic Skin Response*) é um indicador da passagem da corrente eléctrica através da pele e é normalmente medida através de eléctrodos colocados em dois dedos da mão. Uma corrente eléctrica

imperceptível ao corpo humano é aplicada através desses eléctrodos, sendo medida a condutividade entre os eléctrodos.

É conhecido há décadas que a magnitude da resistência à passagem da corrente eléctrica pela pele é afectada não apenas pelo humor do indivíduo mas também pelas reacções emocionais desencadeadas. Embora estes factos sejam conhecidos há mais de 100 anos (o primeiro estudo foi apresentado por Tarchanoff em 1890) apenas nos últimos 25 anos se conheceram as causas que levam a alterações da condutividade da pele, como por exemplo a actividade do sistema nervoso autónomo (Othmer, 2004).

Na última metade do século 19, os investigadores descobriram que a resistência cutânea variava ao longo de todo o corpo. Embora inicialmente esta variação tenha sido descrita como um artefacto, com o aprofundamento das investigações reconheceu-se que a condutividade eléctrica da pele variava sistematicamente com o estado de excitação fisiológica e com os processos mentais. Estas alterações foram descobertas através da actividade das glândulas sudoríparas. Mais tarde, o investigador Carl Jung estabeleceu o GSR como uma medição objectiva da excitação fisiológica e estabeleceu os fundamentos para futuros desenvolvimentos neste campo (Shepherd, 2001).

Facilmente medida e relativamente fiável, o GSR tem sido utilizado como um índice quando se necessita de um parâmetro relacionado com o estado emotivo de um indivíduo. Na Fisiologia, o GSR reflecte o suor gerado pela actividade das glândulas sudoríparas e mudanças no sistema nervoso simpático (o sistema nervoso autónomo é formado pelo simpático, responsável pela activação corporal e reacção a situações, e pelo parassimpático, responsável pelo estado de repouso). A actividade das glândulas sudoríparas em resposta à estimulação do sistema nervoso simpático resulta num aumento do nível da condutividade da pele (Bio-Medical, 2004).

Fisicamente o GSR é uma alteração das propriedades eléctricas da pele em resposta a diferentes tipos de estímulos. Os estímulos tipicamente utilizados são choques eléctricos aplicados a um nervo periférico ou estímulos auditivos. Contudo, qualquer estímulo capaz de provocar um estado de excitação no indivíduo pode ser utilizado. A resposta provocada por esse estímulo é mais dependente da surpresa do estímulo do que a força do próprio estímulo (Uku, 2004).

Existem duas técnicas para a medição da actividade eléctrica da pele. Numa das técnicas uma corrente eléctrica externa é passada através da pele e a resistência à sua passagem é medida; na outra técnica a corrente eléctrica externa não é utilizada, sendo a pele a própria fonte da corrente eléctrica. Através da captação do sinal pelos eléctrodos é possível realizar dois tipos de medições: 1) a condutividade da corrente eléctrica em situações de repouso, e 2) a resposta a um estímulo súbito (como por exemplo um susto) (Othmer, 2004).

É preciso salientar que, na captação do sinal GSR, a epiderme e as glândulas sudoríparas não são estruturas estáticas, isto é, estão a ser dinamicamente reguladas pelos estímulos do meio ambiente, tal como a temperatura e a própria humidade do ar (Rodrigues, 1989). De igual modo, o sinal GSR varia de acordo com o sexo do indivíduo, a alimentação, o tipo de pele e a situação onde se encontra. Por último, a habituação da resposta durante estímulos repetidos também necessita de ser tida em conta.

A actividade eléctrica da pele é caracterizada por alterações lentas (quando comparadas por exemplo com as alterações do batimento cardíaco que são praticamente instantâneas), alterações essas provocadas pelos estímulos da situação que lhe deram origem. Adicionalmente, e após atingidos os valores máximos provocados por esse estímulo, o regresso ao valor basal do GSR é ainda mais lento do que o tempo necessário para se atingirem os valores máximos de GSR (o valor basal é o nível médio de GSR antes da aplicação de um estímulo, tratando-se de uma referência da situação de repouso para comparar com a de estímulo).

Quando um indivíduo se encontra num estado de alarme ou de ansiedade, a condutividade eléctrica da pele aumenta rapidamente em virtude do aumento de actividade das glândulas sudoríparas. Este aumento de actividade aumenta a capacidade de condução da corrente eléctrica pela pele. Nos casos gerais, uma resposta ao stress é acompanhada por um aumento da actividade das glândulas sudoríparas (dai as mãos suadas e frias em situações de stress), em oposição a mãos mais quentes e secas em situações de relaxamento (Othmer, 2004).

Na Figura 2.8 apresenta-se um exemplo de um registo de GSR.

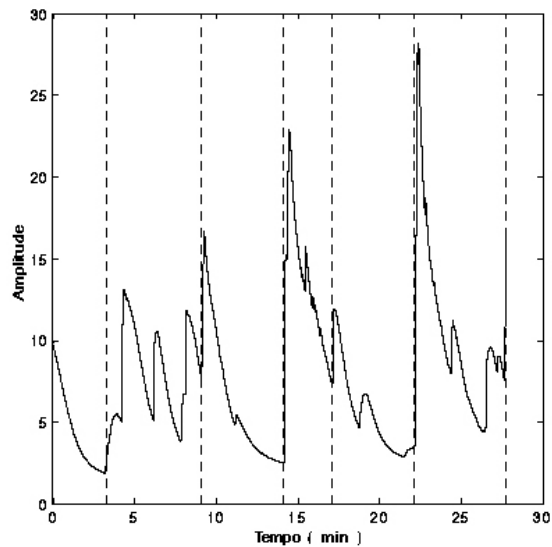


Figura 2.8 – Exemplo de um registo de GSR (Psychology, 2004)

### 2.4.3 Actividade muscular

A electromiografia é uma técnica de registo dos sinais eléctricos associados à actividade muscular de um indivíduo. Para a geração de força muscular pequenas correntes eléctricas são geradas previamente pelas fibras musculares através da troca de iões entre as membranas fibrosas dos músculos. O sinal eléctrico resultante dessa troca de iões pode ser captado, sendo o registo derivado dessa captação denominado de Electromiograma (EMG).

O sinal eléctrico pode ser observado através da aplicação de sensores com características condutivas (eléctrodos) na superfície cutânea ou através de métodos invasivos no músculo (como por exemplo através de agulhas). O EMG superficial, embora menos preciso que o obtido por métodos invasivos, é o mais utilizado uma vez que se trata de uma captação não invasiva e pode ser efectuada para além de pessoal médico especializado com o mínimo risco para a pessoa alvo dessa captação (Day, 2002).

A medição do EMG depende de um variado número de factores e a amplitude do sinal eléctrico pode variar desde  $\mu V$  até valores baixos de mV. As características do sinal captado são dependentes de factores como:

- o tempo de contracção e a intensidade dessa mesma contracção;
- a distância dos eléctrodos ao músculo sobre o qual se faz a recolha do sinal eléctrico;
- as propriedades do tecido envolvente (espessura da pele e do tecido adiposo);
- as próprias propriedades dos eléctrodos e dos dispositivos de medição;
- a qualidade do contacto entre os eléctrodo e a pele (Day, 2002).

Estes factores provocam uma grande variabilidade em termos de registo do EMG. Uma forma de se minimizar essa variabilidade é através da utilização dos mesmos eléctrodos e dispositivos de medição em todos os registos e através do cuidado em assegurar um bom contacto entre os eléctrodos e a pele. Quando se fazem medições em vários indivíduos, uma das formas de minimização da variabilidade do sinal é através da aplicação dos eléctrodos na mesma localização no corpo de todos os indivíduos (Day, 2002).

A qualidade do EMG captado é muitas vezes descrita como um compromisso entre o sinal EMG captado e a captação de ruído indesejado e causado pelo meio ambiente. O objectivo é assim a maximização do sinal EMG e a minimização do ruído. O ruído captado é tipicamente causado por:

- ruído ambiente – ruído gerado por fontes electromagnéticas tais como computadores e dispositivos eléctricos em funcionamento. Basicamente, qualquer dispositivo eléctrico que se encontre ligado à corrente eléctrica é uma fonte de ruído;
- ruído causado pelos dispositivos electrónicos de recolha e registo de sinal – todos os componentes electrónicos produzem algum ruído eléctrico. Este ruído não pode ser eliminado, podendo no entanto ser reduzido através da utilização de componentes electrónicos de alta qualidade e através da utilização de dispositivos melhor desenhados;
- artefactos causados pelo movimento – estes artefactos são causados pelo contacto entre a pele e os eléctrodos

utilizados para a captação dos sinais e pelo próprio movimento do cabo do eléctrodo (De Luca, 2002).

Um EMG que registe uma contracção muscular consiste numa série de descargas do tecido muscular que se encontra na área de captação dos eléctrodos, traduzindo o registo dessa actividade num conjunto de ondas cuja frequência varia desde 20 Hz a 1000 Hz, situando-se a sua amplitude entre 100 e 1000  $\mu\text{V}$ , dependendo estes valores, naturalmente, da quantidade de massa muscular que se encontre associada aos eléctrodos e do grau de contracção do músculo em causa.

Na Figura 2.9 apresenta-se um exemplo de um registo EMG.

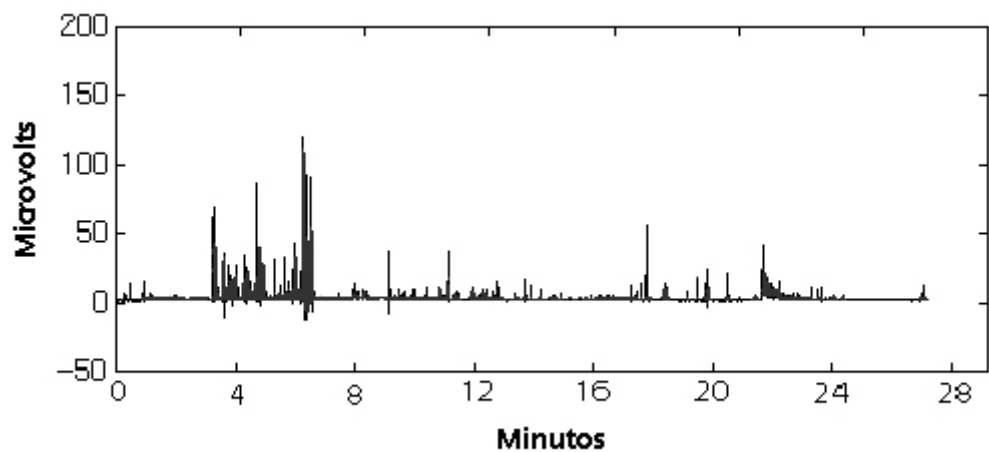


Figura 2.9 – Exemplo de um registo de EMG (MIT, 2004c)

### 3 Ferramentas

O protótipo a ser implementado será composto por vários módulos, cada um desempenhando funções específicas e interligando-se entre si:

- um módulo de aquisição dos sinais fisiológicos (que incluirá o software e o *hardware* necessários);
- um módulo que fará a apresentação em termos visuais e auditivos dos diferentes estados emotivos do utilizador;
- um módulo que possa efectuar a análise dos dados fisiológicos, extraíndo informação sobre esses mesmos dados e obtendo um indicador sobre o estado emotivo;
- módulo de histórico do sistema, para registo dos estados emotivos detectados numa base de dados.

Cada um destes módulos utilizará ferramentas específicas para a sua implementação. Estas ferramentas para os vários módulos, assim como outras alternativas, serão apresentadas a seguir.

#### 3.1 PD – *Pure Data*

O ambiente de programação PD — *Pure Data* — foi desenvolvido por Miller Puckette (membro do Departamento de Música da Universidade da Califórnia, San Diego, e Director Associado do Center for Research in Computing and the Arts), em 1990, para processamento de áudio e manipulação de MIDI — *Musical Instrument Digital Interface* — em tempo real, bem como para trabalhos multimédia interactivos (PD, 2003).

É um projecto *open source* e *freeware*, podendo crescer com a adição de novas extensões. É um código modular com uma interface de funções bem definidas para se integrar com os outros componentes da linguagem, baseados em *externals* (extensões ao programa base) e objectos. Actualmente, são da responsabilidade do IEM — *Institute of Electronic Music and Acoustics* — os instrumentos de divulgação, apoio e partilha de informação sobre o PD, no formato de um portal na *web*, com vários elementos: *mailing lists*, fóruns, uma base de dados de *externals*, documentos, publicações, eventos, comentários e relatórios sobre o PD (IEM, 2003).



O PD é muito semelhante ao programa Max que é um ambiente de desenvolvimento gráfico para música e multimédia, construído por Miller Puckette quando se encontrava a trabalhar no IRCAM — *Institut de Recherche et Coordination Acoustique/Musique* —, em 1980, e mantido pela empresa de software, sediada em San Francisco, nos Estados Unidos, a Cycling74 (Cycling74, 2004b).

### 3.1.1 Introdução

O PD apresenta-se como um ambiente gráfico e com objectos gráficos, que se interligam utilizando o rato e estabelecendo ligações — *patch cords* — entre os *inlets* (entradas) e *outlets* (saídas) dos objectos. Utilizando extensões (*externals*) é também possível fazer processamento e manipulação de vídeo em tempo-real ou diferido.

Para construir um programa em PD existe um conjunto de objectos gráficos (mensagens, objectos, símbolos, comentários, números, gráficos e *arrays*) e um conjunto de objectos utilitários (*sliders* verticais e horizontais, botões de *toggle* e de *bang*, caixas de números, escalas de decibéis e telas de fundo para imagens ou vídeos). Interligam-se por caminhos de ligação (*patch cords*), permitindo que os sinais fluam entre os objectos e de acordo com a sua funcionalidade. Em PD cada programa ou subprograma chama-se *patch*. Por questão de facilidade de desenvolvimento e de depuração de erros, os programas poderão ser subdivididos em subprogramas ou procedimentos (*patches*). Na Figura 3.1 podem ser vistos dois exemplos de *patches* em PD.

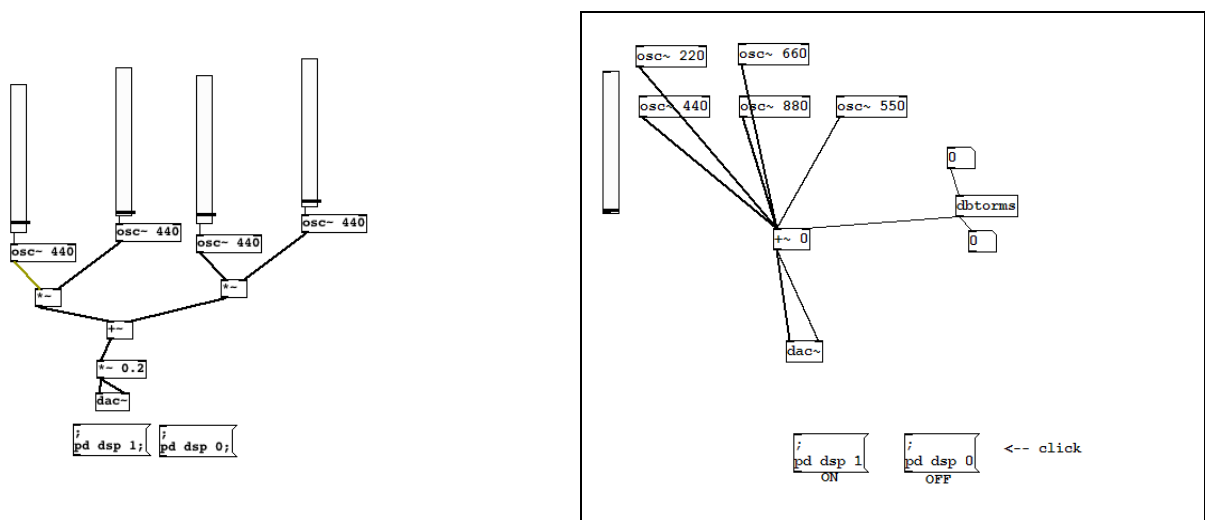


Figura 3.1 – Exemplo de dois *patches* em PD

O PD, sendo *freeware*, está acessível a todos, sendo possível acrescentar-lhe novas facilidades, desenvolvidas em C ou C++, adicionadas como bibliotecas ou *externals* novos ou simplesmente novos objectos. A diferença fundamental existente entre *externals* e bibliotecas é que nas bibliotecas os procedimentos que implementam novas funções estão agrupados por funcionalidade, enquanto que nos *externals* poderão estar dispersos por vários procedimentos ou rotinas.

### **3.1.2 GEM – *Graphics Environment for Multimedia***

GEM, *Graphics Environment for Multimedia*, é uma extensão das bibliotecas do PD que possibilitam a criação de gráficos 3D OpenGL e a manipulação de imagens, texto e vídeo em tempo-real. Foi originalmente concebido por Mark Danks para a geração de gráficos no computador em tempo-real, sendo especialmente vocacionado para composições audiovisuais e é agora mantido por IOhannes M Zmölning do *Institute of Electronic Music and Acoustics*. Pode correr em ambientes Windows NT 4.0 ou posteriores, Linux, Macintosh e SGI Irix 6.2 ou posterior (GEM, 2003).

O GEM não é uma aplicação, tratando-se de uma biblioteca que o PD carrega como uma extensão das suas capacidades básicas. Uma vez instalado o GEM, todos os objectos incluídos nesta biblioteca são tratados da mesma forma que os objectos originais do PD, ou seja, os objectos do GEM são criados da mesma forma que os do PD, possuindo de igual modo *inlets* e *outlets*, argumentos de criação, etc.

Existem diversos objectos destinados a funções diferentes, desde objectos para criação de gráficos poligonais, iluminação, mapeamento de texturas, processamento de imagem, até objectos para processamento de vídeo e de entradas de câmaras de vídeo ou *web cams*, etc. Tudo é possível em tempo-real, sendo visível a combinação de áudio e gráficos, controlando um tipo de média a partir de outro.

### **3.1.3 Lista de objectos do GEM**

Todos os objectos do GEM encontram-se dentro de um dos grupos a seguir apresentados, agrupados de acordo com a sua função:

- *Controls* - Estes objectos são usados em todos os *patches* que necessitem de funcionalidades do GEM, sendo por isso considerados os seus objectos base. São estes objectos que controlam os gráficos e todo o mecanismo de *rendering* em tempo real;
- *Geos* - Estes são os objectos que possibilitam a criação de formas visuais pré-definidas, mas que podem ser alteradas à medida das necessidades. Estas formas incluem cubos, círculos, quadrados, etc.;
- *Nongeos* - Objectos do GEM que não possuem uma forma explícita, mas que afectam o processo de *rendering* de alguma forma. Exemplo de objectos *Nongeos*: *light*, que cria um ponto de luz, *world\_light*, que define uma iluminação global;
- *Manipulators* - São objectos destinados à manipulação de qualquer objecto *Geos* presente numa janela do GEM. Esta manipulação pode passar, por exemplo, pela rotação de uma imagem, pela alteração da dimensão de um vídeo ou por uma translação no eixo XYZ desse mesmo vídeo;
- *Particles* - Objectos do GEM que envolvem o sistema de partículas. Este sistema pode ser usado para a criação de efeitos de fogo, fumo, água, etc.;
- *Pixes* - Estes são objectos de processamento de imagem e vídeo do GEM. O número de objectos desta categoria é bastante elevado, incluindo objectos para a manipulação do nível de alfa (*alpha*) de imagens (nível de transparência) e vídeos, adição, subtracção e multiplicação de imagens e vídeos, etc.;
- *openGL* - Objectos destinados à criação de gráficos OpenGL. Cada um destes objectos (mais de 250) possui o prefixo GEM seguido do nome do objecto (por exemplo, [GEMglVertex3f]);
- *TV* - Objectos destinados ao processamento de imagens baseados no domínio temporal. Por exemplo, o objecto *tv\_movement* altera o canal *alpha* de um vídeo baseado na diferença entre duas imagens anteriores;

- *MarkEx* - Uma colecção de objectos genéricos para a manipulação de dados, especialmente indicados para o uso com outros objectos do GEM.

### 3.1.4 Aplicações do PD/GEM

O PD, por se tratar de ambiente de programação visual orientado para a manipulação de áudio e, com biblioteca GEM, a manipulação de vídeo, imagens e texto em tempo-real, é um sistema de criação multimédia bastante flexível. Por se tratar também de um ambiente de programação extensível com novos objectos desenvolvidos em C/C++, as suas aplicações são apenas limitadas pela criatividade do utilizador/programador. Apresentam-se de seguida, alguns exemplos de aplicações do PD:

- criação de música interactiva – o PD possui vários objectos específicos para a análise e o processamento de informação MIDI para a criação musical. Além disso, como essa informação MIDI é numérica e os valores variam de entre 0 e 127, todos os objectos que lidam com números podem também ser utilizados para analisar e processar estes valores. Um exemplo de uma aplicação MIDI é a de acompanhamento em tempo real dos dados MIDI recebidos pelo PD. Um simples teclado musical com saída MIDI pode enviar essa informação relativa às notas tocadas. Analisando e processando essas notas no PD é possível criar, por exemplo, acordes que acompanham as linhas melódicas produzidas no teclado MIDI;
- emulador de sintetizadores em *hardware* – as funcionalidades de síntese de áudio são bastante extensas no PD. Possui todos os componentes necessários para a criação de um sintetizador em software análogo aos sintetizadores em *hardware*. Esses componentes incluem osciladores, filtros, amplificadores, etc. Com os conhecimentos de síntese necessários, é possível criar uma réplica fiel de um dos sintetizadores mais populares, como por exemplo o Yamaha DX7. Além da emulação de sintetizadores já existentes, é também possível criar sintetizadores com sonoridades completamente novas e originais;

- instalações multimédia – uma outra aplicação do PD é a da sua integração em instalações multimédia. A facilidade de integração de áudio, vídeo, imagem e texto, aliada à possibilidade de adição de sensores em *hardware* externos que podem comunicar com o PD e controlar em tempo real toda a componente visual e sonora, tornam-no uma ferramenta poderosa para a criação de instalações multimédia;
- criação de áudio, imagem e vídeo experimental – o PD pode ser utilizado para a criação de material sonoro e visual de carácter experimental. É possível por exemplo explorar as capacidades do áudio na criação visual e vice-versa: “o som pode ser visto e as imagens e vídeos podem ser ouvidos”;
- criação musical colaborativa – o PD permite a criação de programas distribuídos por vários computadores de modo a que todos os utilizadores possam contribuir para a criação musical. Um computador com o PD instalado pode ser utilizado como sendo o servidor da criação áudio. Outros computadores poderão funcionar como controlos remotos desse servidor. Através de protocolos como TCP (*Transport Control Protocol*), UDP (*User Datagram Protocol*), MIDI ou o recente OSC (*Open Sound Control*), é possível sincronizar vários patches PD presentes em vários computadores de forma a que um utilizador possa alterar em tempo real a produção musical no seu patch, sendo essa alteração difundida para todos os outros utilizadores que se encontrem a utilizar o sistema. Assim, qualquer alteração no som efectuada por um utilizador é imediatamente ouvida por qualquer outro utilizador;
- interligação e interacção entre o PD e outras tecnologias – é possível interligar programas em PD com outras linguagens e ferramentas tais como Java, Bases de Dados, Flash (através do Flashserver), etc.

### 3.2 Max/MSP e Jitter

Max/MSP é um ambiente de programação visual para o processamento de MIDI, áudio e multimédia. Originalmente disponível apenas para o sistema operativo Macintosh, encontra-se

actualmente disponível também para Windows XP (Cycling74, 2004b).

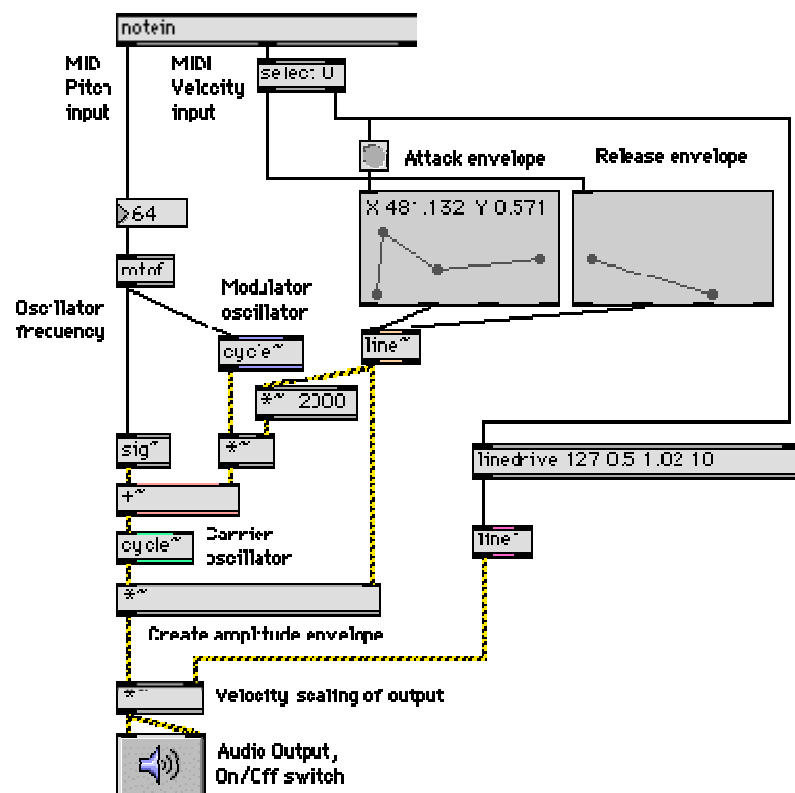


Figura 3.2 – Max (Cycling74, 2004b)

Tratando-se tal como o PD de um ambiente de programação visual, é possível criar aplicações visualmente através de um conjunto de objectos que se ligam através de *patch cords*.

O ambiente básico que inclui objectos de processamento de MIDI, controlo, interface e de temporização/sincronismo é denominado Max. Existem no entanto 2 bibliotecas de objectos que aumentam as suas funcionalidades básicas:

- **MSP** - conjunto de objectos para o processamento de áudio, desde filtros e síntese áudio, até à gravação em disco duro de áudio (Cycling74, 2004c);

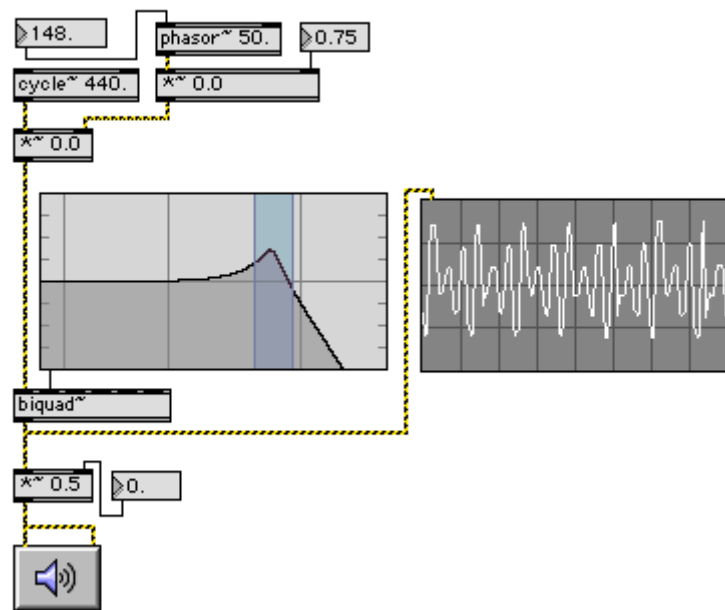


Figura 3.3 – MSP

- **Jitter** - conjunto de 135 objectos para o processamento de vídeo, matrizes e gráficos 3D. Estes objectos ampliam as funcionalidades do Max/MSP através da manipulação de dados em matrizes (quaisquer tipos de dados que possam ser representados em linhas e colunas, tais como vídeo e imagens, geometria 3D), assim como texto ou sistemas de partículas 3D (Cycling74, 2004a).

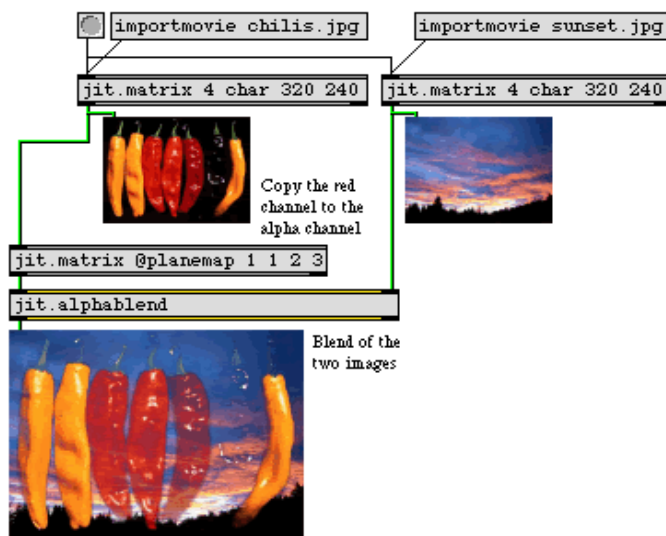


Figura 3.4 – Jitter

### **3.3 MIDI – *Musical Instrument Digital Interface***

MIDI, *Musical Instrument Digital Interface*, é uma norma definida por várias empresas proeminentes no campo da construção de instrumentos musicais electrónicos, para a transferência de informação entre esses instrumentos, inclusive entre computadores equipados com MIDI (Moog, 2004).

Enquanto que os cabos de áudio permitem a transferência de som entre instrumentos, o MIDI proporciona a transferência de comandos musicais entre dois ou mais equipamentos equipados com MIDI. Por exemplo, um músico pode tocar dois ou mais dispositivos sonoros (tais como sintetizadores) através de um único teclado MIDI, sincronizar uma rede de instrumentos MIDI através de um sinal de sincronismo MIDI, ou utilizar um computador para gravar, processar e reproduzir novamente a informação MIDI recebida.

Tecnicamente, o MIDI é uma interface digital série e bidireccional. O termo “digital” significa que toda a informação transmitida é representada como um conjunto de números. Por exemplo, quando num teclado MIDI se pressiona uma tecla, o número a que corresponde essa tecla é transmitido. O termo “série” significa que esses números são transmitidos um bit de cada vez. Segundo a especificação MIDI, a taxa de transferência de informação é de cerca de 30,000 bits por segundo. Um comando típico MIDI, que normalmente possui cerca de 13 bits, demora cerca de 0.3 milésimos de segundo a transmitir. O termo “bidireccional” significa que a informação pode circular nos dois sentidos entre dois equipamentos MIDI. Por exemplo, um teclado MIDI pode estar ligado a um computador equipado com uma interface MIDI e este pode gravar e reproduzir a informação MIDI.

Embora o protocolo MIDI tenha sido desenvolvido originalmente tendo em vista a comunicação entre instrumentos musicais electrónicos e computadores, pode também ser utilizado para a transmissão de uma grande variedade de sinais de controlo não limitados apenas ao campo musical. Como exemplo, o MIDI é frequentemente utilizado para o controlo de luzes de espectáculos, no envio de informação para movimento de robôs, etc. (Moog, 2004).



### 3.4 WaveRider e WaveWare

O WaveRider é o nome de uma gama de sistemas de monitorização para o ambiente Windows que capta informação relativa a ondas cerebrais e outros sinais biológicos, tais como batimento cardíaco (HR), condutividade eléctrica da pele (GSR) e sinais musculares (EMG). Todos os sinais podem ser visualizados em tempo-real (MindPeak, 2003a).

O WaveRider possui uma série de funções pré-programadas destinadas à utilização comum do sistema (por exemplo, uma função destinada à leitura das ondas cerebrais). Estas funções podem no entanto ser alteradas de acordo com as necessidades dos utilizadores. Além disso, o software do WaveRider possui algum nível de customização, possibilitando mostrar vários sinais, analisá-los e convertê-los em MIDI (MindPeak, 2003a).

Na gama WaveRider existem 2 produtos, o WaveRider Pro e o WaveRider Jr. O WaveRider Pro possui 4 canais de utilização diversa e um canal adicional dedicado à captação do sinal GSR. O WaveRider Jr possui 2 canais de utilização diversa e um outro canal dedicado à captação do sinal GSR (MindPeak, 2003a).

O WaveRider é composto pelos eléctrodos, o *hardware* e o *software* de suporte (WaveWare). Os eléctrodos captam os sinais biológicos e são afixados ao corpo e ligados ao *hardware*. O *hardware* transmite os sinais para o computador através de um cabo série. O *software* Windows do WaveRider, o WaveWare, processa, analisa e permite a visualização dos sinais.

O *hardware* do WaveRider possui um canal dedicado à captação do sinal GSR e vários canais (dependendo do modelo) de utilização diversa. Estes canais, com baixo nível de ruído, podem captar sinais cerebrais, musculares e cardíacos. Cada um destes canais de utilização diversa pode ser configurado para captar qualquer sinal biológico (exceptuando o sinal GSR para o qual há um canal dedicado).

O WaveWare é o *software* Windows utilizado com o WaveRider. O *software* apresenta várias janelas com gráficos para os sinais captados. Fazendo duplo clique em qualquer dessas janelas abre-se uma caixa de diálogo que controla o processamento desse sinal e as características dessa janela.

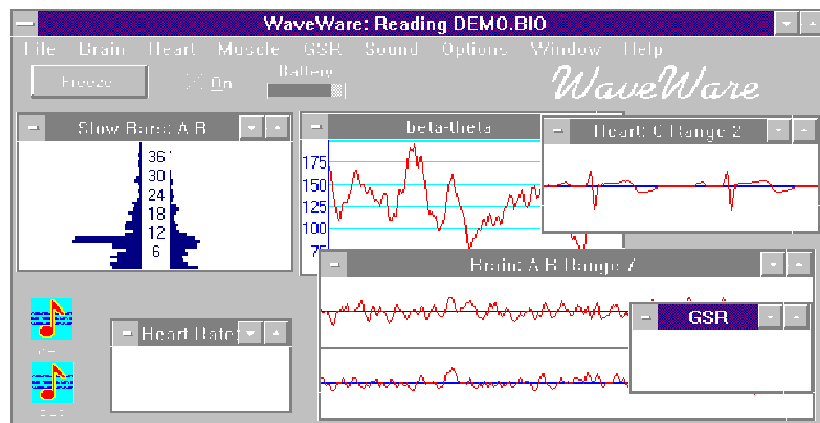


Figura 3.5 – WaveWare (MindPeak, 2003a)

É possível monitorizar, guardar e analisar sinais biológicos. Os sinais capturados podem ser exportados para posterior análise noutros programas ou para interacção em tempo real com aplicações desenvolvidas para o efeito.

### 3.5 Outros sistemas de captação de sinais fisiológicos

#### 3.5.1 Procomp+/Biograph

O ProComp+ é um sistema multi modo, de 8 canais, para a monitorização de EMG, actividade das ondas cerebrais (EEG – *Electroencephalogram*), temperatura do corpo, HR, BVP, GSR e respiração. É possível monitorar apenas um sinal ou qualquer combinação desses sinais, de acordo com as necessidades específicas (Bio-Medical, 2003).

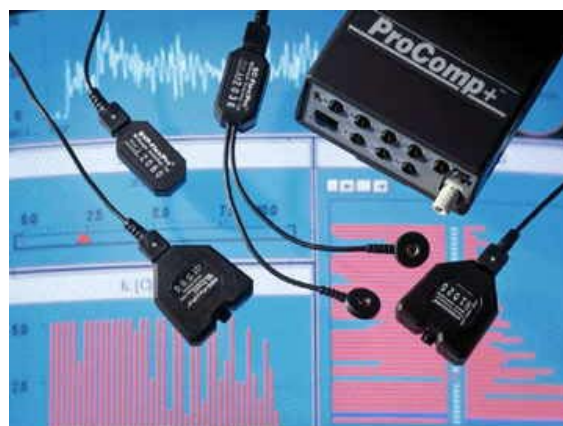


Figura 3.6 – ProComp + (Bio-Medical, 2003)

### 3.5.2 BioMuse

O BioMuse, apresentado pela empresa BioControl Systems em 1992, é utilizado como uma plataforma de desenvolvimento para uma diversidade de aplicações. É basicamente um controlador com 8 canais de entrada para sinais bioeléctricos, que possibilita aos utilizadores o controlo de funções no computador directamente a partir dos músculos, movimento ocular ou ondas cerebrais, evitando completamente o uso de sistemas de entrada habituais como o rato ou o teclado (BioControl, 2003).

O sistema recebe sinais das quatro principais fontes de actividade eléctrica no corpo humano: músculos (EMG), movimento ocular (EOG - *Electro-Oculography*), coração (ECG) e ondas cerebrais (EEG). Estes sinais são adquiridos utilizando eléctrodos não evasivos.

O BioMuse amplifica e digitaliza o sinal biológico, processando-o de seguida com uma biblioteca de algoritmos de processamento de sinal própria da BioControl. Para uma dada aplicação, o algoritmo apropriado tem como saída código que pode controlar qualquer dispositivo com uma interface digital.



Figura 3.7 – BioMuse (BioControl, 2003)

### 3.5.3 IBVA – *Interactive Brainwave Visual Analyser*

O sistema IBVA (*Interactive Brainwave Visual Analyser*), da IBVA Technologies, Inc, é um sistema que capta a actividade das ondas cerebrais em tempo real e que permite usar os sinais captados para controlar música, filmes, imagens, sons, outro software, ou qualquer

outro dispositivo electrónico através de um computador local ou através da Internet (IBVA, 2003).

O utilizador não precisa de estar fisicamente ligado ao computador: uma banda para a cabeça, utilizada para a captação da actividade das ondas cerebrais, pode comunicar com um receptor num raio de 30 pés, permitindo deste modo um elevado grau de liberdade ao utilizador.

O software fornecido com o IBVA pode comunicar os dados relativos às ondas cerebrais a qualquer outro computador, em tempo real, através de um simples modem ou de uma ligação em rede ou à Internet

O *software* possui janelas para a representação gráfica das formas de onda relativas às ondas cerebrais. Além dessa representação, possui facilidades para a ligação e transmissão desses dados a qualquer outro computador numa rede e para a conversão da informação cerebral em MIDI, entre outras funções (IBVA, 2003).

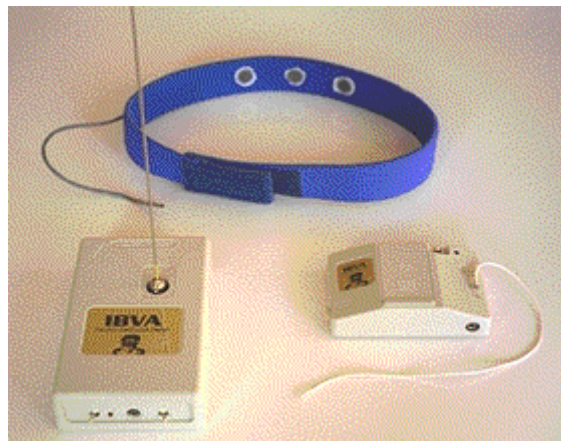


Figura 3.8 – Sistema IBVA (IBVA, 2003)

### 3.6 MIDI YOKE - MIDI *Patch Cable driver*

O MIDI Yoke é um programa utilizado para a ligação de qualquer saída de um programa MIDI à entrada de um outro programa MIDI. A informação MIDI é assim passada directamente da saída de um programa para a entrada de um outro (MIDI/OX, 2003). O esquema seguinte pretende mostrar um exemplo de utilização:

**[ Sequencer 1 Out ]==>[ Out MIDI Yoke In ]==>[ In Sequencer 2 ]**

O MIDI Yoke pode ser configurado de forma a disponibilizar um número variado de portas MIDI, entre 1 e 16. Cada uma dessas portas pode ser utilizada para estabelecer a comunicação entre dois ou mais programas (no máximo de 4 programas para cada porta). De acordo com o protocolo MIDI, cada porta MIDI possui 16 canais, de forma que podemos estabelecer a ligação entre 2 programas com 16 canais, tendo cada um desses canais informação MIDI independente da dos restantes.

Na ligação entre duas aplicações MIDI, a primeira aplicação define a porta MIDI Yoke 1 como a porta de saída, enquanto que a segunda aplicação define a porta MIDI Yoke 1 como a porta de entrada. Os dados MIDI passarão assim da primeira aplicação MIDI para a segunda, através do MIDI Yoke (MIDIOX, 2003).

### 3.7 Java

A linguagem Java é uma linguagem de programação orientada a objectos, criada por James Gosling e outros engenheiros da empresa Sun Microsystems. Foi desenvolvida em 1991, como parte do projecto «Green Project» e foi oficialmente disponibilizado em 1995 pela SunWorld. Inicialmente, esta linguagem chamou-se Oak, em homenagem a uma árvore (carvalho) existente perto do escritório dos criadores (Java, 2004a).

No Java, salientam-se alguns aspectos que o caracterizam e justificam a sua vasta divulgação e utilização como plataforma de desenvolvimento. É uma linguagem de programação orientada a objectos (OO), o que significa que o projecto do software decorre em torno de “módulos” (ou seja, objectos). O Java pode funcionar como uma linguagem de programação, como *middleware* e como uma plataforma; utiliza para isso a Java Virtual Machine (JVM), que permite ser independente da plataforma onde corre.

O código escrito em Java é executado numa máquina virtual (VM – *Virtual Machine*) onde os programas são traduzidos para uma linguagem intermédia, em *bytecodes*, que, por sua vez, são interpretados pelo processador em vez de serem directamente convertidos para código máquina. Assim, o Java tornou-se uma linguagem portátil e independente da plataforma onde corre. Contudo, o grande número de plataformas de desenvolvimento

diferentes leva a que a expressão «*Write once, run anywhere*» da Sun se torne «*Write one, debug everywhere*».

Mesmo assim, é vastamente utilizado pelos seus recursos, tais como: recursos de rede, possuindo uma biblioteca de funções que permitem trabalhar com os protocolos TCP/IP (*Transmission Control Protocol/Internet Protocol*), com o HTTP (*Hypertext Transfer Protocol*) e o FTP (*File Transfer Protocol*); aspectos de segurança implementados na linguagem, podendo executar programas via rede com restrições de segurança; características da linguagem de programação, próxima da linguagem C, com um grande conjunto de bibliotecas (API's); simplicidade na especificação quer da linguagem quer do ambiente de execução (JVM); e características de gestão dos recursos da máquina, como sejam a deslocação automática da memória através de *garbage collection* e o carregamento dinâmico do código, que se traduz no carregamento das classes Java à medida que são necessárias.

O Java apresenta as seguintes áreas: J2SE (*CoreDesktop*), J2EE (*Enterprise Server*), J2ME (*Mobile/Wireless*), Java Card, Java Web Services, XML, e ainda, outras tecnologias Java (Java, 2004a).

### 3.8 Matlab

Matlab resulta etimologicamente de «*matrix laboratory*». Foi desenvolvido inicialmente, em 1970, por Cleve Moler, a trabalhar na Universidade do Novo México, para aceder ao software, escrito em Fortran, para a resolução de problemas de álgebra linear numérica (LINPACK) e de valores próprios (EISPACK). Em 1980 o Matlab foi reescrito em C e com mais funcionalidades. Só em 1984, a empresa MathWorks, Inc. começaria a comercializar o Matlab e a ser responsável pelo desenvolvimento, vendas e apoio técnico. É um *package* poderoso para computação técnica, integrando computação, visualização e programação, utilizando a notação matemática quer para integrar e processar os dados, quer para devolver os resultados (MathWorks, 2004a).

O Matlab, sendo um ambiente de desenvolvimento interpretado, pode funcionar em tempo real em modo interpretado, executando funções pré-construídas pelo utilizador ou do Matlab, e como uma linguagem de programação, compilada, existindo tabelas de compatibilidade que indicam quais as funcionalidades e facilidades

que podem ser compiladas ou não. Como foi desenvolvido inicialmente para computação científica, não é adequado, por exemplo, para efectuar o *parsing* de texto.

As suas áreas de aplicação são: a computação técnica, isto é, computação matemática, análise, visualização e desenvolvimento de algoritmos; projecto de controlo, isto é, projecto baseado num modelo de controlo de sistemas, incluindo simulação, prototipagem rápida e geração de código para sistemas embebidos; processamento de sinal e comunicações, isto é, projecto baseado em modelos de processamento de sinal e de sistemas de comunicação, incluindo simulação, geração de código e verificação; processamento de imagem, isto é, aquisição de imagem, análise, visualização e desenvolvimento de algoritmos; teste e medições, traduzindo-se em conectividade de *hardware* e análise de dados para aplicações de teste e de medição; e, por último, modelação financeira e análise, permitindo modelação financeira, análise e aplicações de distribuição (MathWorks, 2004a).

### **3.9 Redes neuronais**

As redes neuronais são um paradigma para o processamento da informação, inspirada na forma como os sistemas nervosos biológicos, tal como o cérebro humano, processam a informação.

Uma rede neuronal é composta por um grande número de elementos de processamento interligados (neurónios), trabalhando em conjunto para a resolução de problemas específicos (MathWorks, 2004b).

As redes neuronais, tal como nos humanos, obtêm o conhecimento através da aprendizagem. Uma rede neuronal é assim configurada para uma aplicação específica, tal como o reconhecimento de padrões ou a classificação de dados, através de um processo de aprendizagem. Essa aprendizagem envolve ajustamentos nas ligações sinápticas entre os neurónios, tal como nos sistemas biológicos.

Embora se pense que as simulações de redes neuronais sejam um conceito recente, o facto é que este já existe mesmo antes da existência de computadores. O primeiro neurónio artificial foi criado por Warren McCulloch e Walter Pitts, embora a tecnologia existente na altura não lhes permitisse grandes avanços neste campo (Loder, 2004).

As redes neuronais, com a sua potencial capacidade de extrair significado de conjuntos de dados complexos ou imprecisos, podem ser usadas para identificação de padrões ou para a detecção de tendências demasiado complexas para serem detectadas por humanos ou outras técnicas computacionais. Uma rede neuronal pode ser pensada como um perito na área de informação na qual foi treinado para analisar.

As vantagens das redes neuronais incluem:

- aprendizagem adaptativa: possuem a capacidade de aprenderem como realizar tarefas com base nos dados iniciais de aprendizagem;
- auto-organização: uma rede neuronal pode criar a sua própria organização ou representação da informação recebida durante o período de aprendizagem;
- funcionamento em tempo real: as operações das redes neuronais podem ser efectuadas em paralelo;
- tolerância a falhas através da eliminação de informação redundante: destruição parcial de ramos da rede com vista a uma melhoria do desempenho não afecta o funcionamento.

As redes neuronais adoptam um conceito diferente na resolução de problemas, quando comparados com os métodos computacionais convencionais. Os métodos tradicionais utilizam uma abordagem predominantemente logarítmica, isto é, seguem um conjunto de instruções para a resolução de problemas. Ora, esta abordagem implica que os passos necessários para a resolução desse problema sejam conhecidos, o que limita a capacidade de resolução de problemas aos que são actualmente conhecidos e já se sabe como resolver.

Uma vez que tentam processar a informação da mesma forma que um cérebro, a rede é composta por um grande número de elementos de processamento (neurónios), altamente interligados entre si e trabalhando em paralelo na resolução de um problema específico. As redes aprendem através de exemplos que deverão ser seleccionados cuidadosamente para que o tempo de aprendizagem não seja desperdiçado ou, pior ainda, para que a rede não funcione sequer. A desvantagem das redes é que, uma vez que estas aprendem sozinhas a resolver determinado problema, o seu comportamento pode ser aparentemente imprevisível



Os sistemas de computação algorítmica e as redes neuronais não se encontram em competição, complementam-se. Existem situações mais vocacionadas para um ou para o outro. Mais ainda, muitas vezes, um grande número de tarefas necessitam de sistemas que utilizem uma combinação dos dois paradigmas para a obtenção dos melhores resultados (por exemplo, um computador convencional pode ser usado para supervisionar o funcionamento da rede neuronal).

### 3.9.1 Modelo do neurónio artificial e arquitectura da rede neuronal

#### 3.9.1.1 Neurónio com entrada simples

Um neurónio artificial, por analogia com o neurónio biológico, possui características de funcionamento tais como capacidade de aprendizagem, memorização e processamento.

A representação de um neurónio artificial é mostrada na Figura 3.9 (Casanova, 2004).

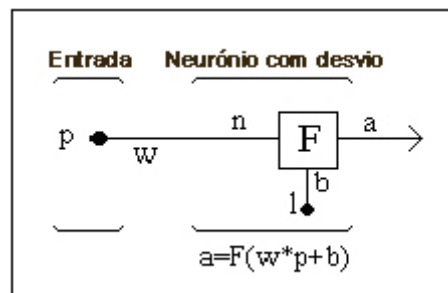


Figura 3.9 – Representação de um neurónio artificial

Este neurónio possui uma entrada simples, com peso  $w$  e desvio (*bias*)  $b$ . O peso e o desvio são dois factores auto-ajustáveis no decorrer do treino da rede neuronal, fazendo-se referência ao treino da rede mais adiante.

O argumento da função de transferência ( $F$ ) é a soma do produto  $w*p$  e do desvio  $b$ . Os pesos são os valores da rede que quando processados conjuntamente com os dados fornecidos ( $p$ ), permitem que a rede possa disponibilizar os valores de saída pretendidos.

O peso é o primeiro factor que vai ser exposto ao estímulo do neurónio (entrada  $p$ ). A função de transferência  $F$  é normalmente

uma função degrau ou uma função sigmoid, podendo no entanto ser utilizadas muitas outras, dependendo da aplicação desejada (Casanova, 2004).

### 3.9.1.2 Exemplos de funções de transferência

Existem várias funções de transferência, das quais se destacam as mais utilizadas:

- Função de transferência Hard Limit - limita a saída do neurónio aos valores 0 e 1;

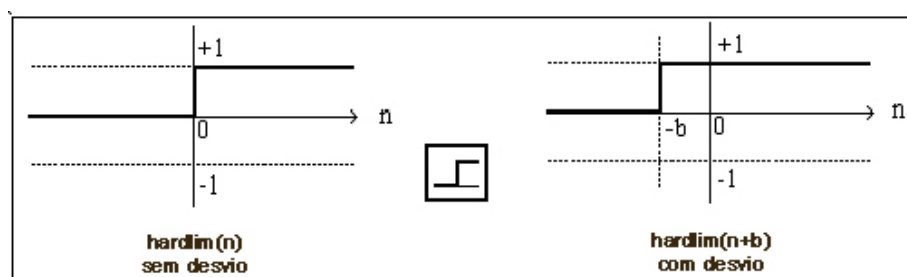


Figura 3.10 – Função de transferência Hard Limit (Casanova, 2004)

- Função de transferência Linear - a saída do neurónio é função linear da entrada;

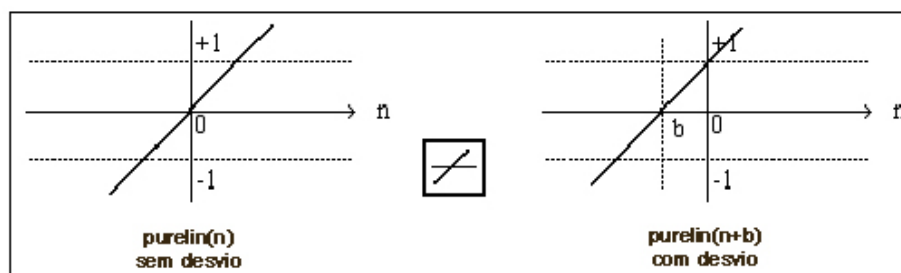


Figura 3.11 – Função de transferência Linear (Casanova, 2004)

- Função de transferência Sigmoid - limita a saída do neurónio ao intervalo de valores entre 0 e 1. É muito utilizada por ser uma função diferenciável.

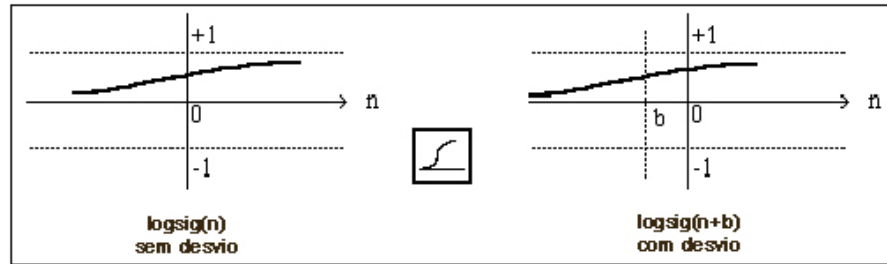


Figura 3.12 – Função de transferência Sigmoid (Casanova, 2004)

### 3.9.1.3 Neurónio com entradas múltiplas

Um único neurónio com R entradas é apresentado na Figura 3.13.

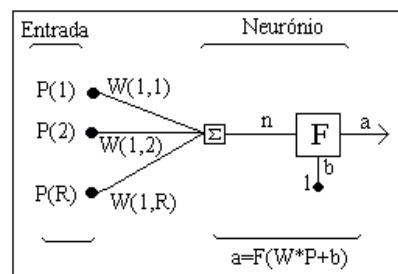


Figura 3.13 – Representação de um neurónio com R entradas (Casanova, 2004)

A matriz de pesos  $W(1,j)$  e a matriz de entradas  $P(j)$  podem ser representadas por um vector linha  $W$  e por um vector coluna  $P$  respectivamente, como a seguir se representa (Casanova, 2004):

$$W = [W(1,1) \quad W(1,2) \quad \dots \quad W(1,R)] \quad P = \begin{bmatrix} P(1) \\ P(2) \\ \vdots \\ P(R) \end{bmatrix}$$

Neste caso entradas individuais  $P(j)$  são pesadas pelos elementos  $W(1,j)$  da matriz  $W$  e somados para formar as entradas pesadas da função de transferência  $F$ . O neurónio tem um desvio  $b$  e uma saída  $a$ .

### 3.9.1.4 Uma camada de neurónios

Uma camada (*layer*) de neurónios com R entradas e S neurónios é mostrada na Figura 3.14.

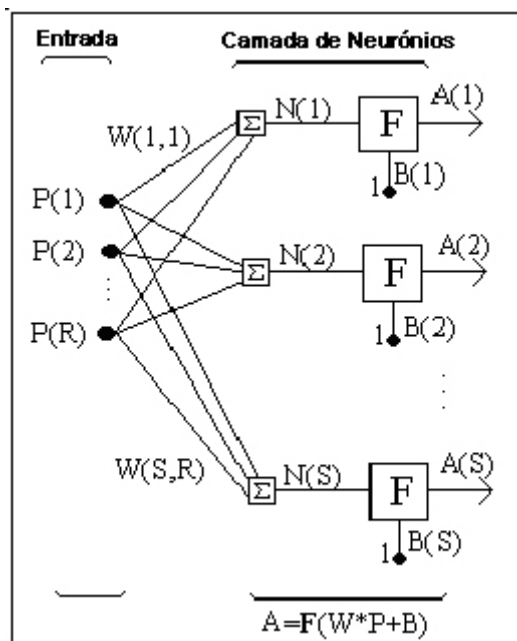


Figura 3.14 – Representação de uma camada de neurónios com R entradas e S neurónios (Casanova, 2004)

Neste caso, cada elemento da entrada é ligado a todos os neurónios de entrada, sendo multiplicado pelo peso correspondente.

Cada neurónio S produz um valor, constituindo esse valor a saída que faz parte de um vector N. O valor na entrada da função de transferência é a soma de produto  $W * P$  e do desvio B, sendo estes elementos da soma os argumentos da função de transferência. Finalmente, a saída da camada de neurónios é dada por um vector coluna A.

O número de entradas numa camada de neurónios pode ser diferente do número de neurónios. Uma camada não é necessariamente construída para ter um número de entradas igual ao número dos seus neurónios.

### 3.9.1.5 Camadas múltiplas de neurónios

Este caso, por ser em tudo semelhante ao caso anterior, omitiu-se aqui a sua descrição. Neste tipo de camadas múltiplas a saída de

uma camada irá ser a entrada da camada seguinte, sendo todas as operações de cálculo idênticas ao caso de uma camada simples.

### 3.9.1.6 Arquitecturas recursivas

As arquitecturas neuronais recursivas são diferentes dos tipos de arquitecturas apresentadas anteriormente, uma vez que as suas saídas são ligadas às entradas. Apresenta-se na Figura 3.15 uma arquitectura recursiva com S neurónios.

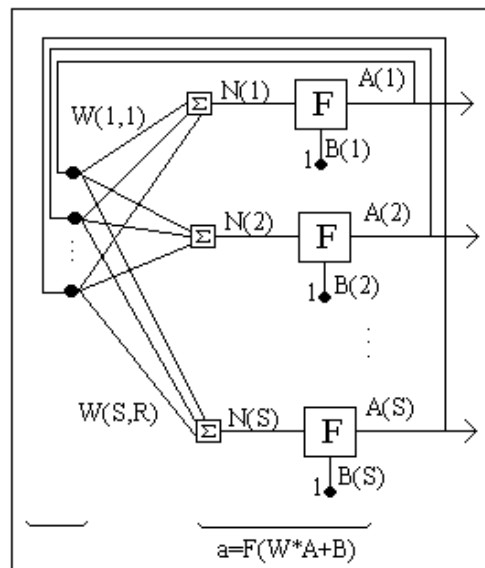


Figura 3.15 – Representação de uma arquitectura neural recursiva com S neurónios (Casanova, 2004)

Aplicando à entrada Q vectores simultaneamente, esta rede neuronal produz como resultado um vector de saída  $A(j)$  para cada vector de entrada  $P(j)$ . Concluindo, a aplicação de Q vectores de entrada produzem uma matriz de saída A com S linhas e Q colunas. Este tipo de rede é sobretudo utilizado em sistemas de retroacção e controlo.

### 3.9.2 Aprendizagem

A capacidade de aprendizagem é o princípio fundamental da inteligência. Embora seja difícil de formular uma definição precisa sobre aprendizagem, um processo de aprendizagem numa rede neuronal pode ser visto como um problema de adaptar a arquitectura e os pesos das ligações neuronais para que a rede possa resolver um problema específico eficientemente (Casanova, 2004).

As redes neuronais devem aprender os pesos das ligações efectuadas nos treinos modelo, sendo capazes de aprender automaticamente com exemplos elucidativos.

Em vez de seguir regras especificadas por humanos, as redes neuronais aprendem regras subjacentes a uma colecção de exemplos representativos. Esta é uma das maiores vantagens das redes neuronais em relação a outros sistemas de inteligência artificial.

### **3.9.3 Métodos de treino de redes neuronais**

Existem vários métodos de treino de redes, que se adaptam mais eficientemente aos problemas para que foram criadas. Podem, no entanto ser utilizadas para outros fins, mas, com menos eficiência. De seguida faz-se uma breve referência a algumas das teorias mais usadas para o treino de redes neuronais.

#### **3.9.3.1 Perceptron**

Esta teoria criada por Rosenblatt baseia-se numa rede com uma só camada de neurónios onde os seus pesos e desvios devem ser treinados para produzir um vector de saída correcto, desde que o vector de entrada seja similar aos vectores onde a rede foi treinada (Casanova, 2004).

A técnica de treino usada é chamada de regra de aprendizagem perceptron. Esta técnica é especialmente utilizada em problemas simples de classificação.

Contudo esta teoria possui várias limitações (daí que só se utilize em problemas simples). Primeiro, a saída está limitada a um de dois valores, pois a função de transferência utilizada é uma função degrau, ou seja, só toma valores 0 ou 1. Segundo, esta teoria só consegue classificar correctamente dados que sejam linearmente separáveis.

#### **3.9.3.2 Regra de Widrow-Hoff Learning**

Esta teoria foi desenvolvida por Widrow e Hoff. O modelo de rede neuronal difere do modelo anterior por possuir uma função de transferência linear, ou seja, pode-se ter na saída quaisquer valores

e não só valores zero e um como no modelo anterior (Casanova, 2004).

Os vectores de entrada são usados para treinar a rede (ajustando os seus pesos e desvios) para produzir o vector de saída correcto.

Esta teoria é muito utilizada em problemas de aproximação de funções lineares e associações tipo.

A arquitectura da rede neuronal é construída completamente com o problema que se pretende resolver, isto é, o número de entradas na rede e o número de neurónios da rede é construído respectivamente pelo número de entradas e pelo número de saídas requeridas pelo problema.

Esta teoria permite a utilização de camadas múltiplas, no entanto a performance da rede não é aumentada com esta solução.

### **3.9.3.3 Backpropagation**

Esta teoria foi criada para generalizar a teoria de Widrow-Hoff para múltiplas camadas de neurónios, funções de transferência não lineares e diferenciáveis (Casanova, 2004).

A determinação do melhor “caminho” a seguir durante o treino é feita à custa da derivada da função associada ao erro, ou seja, para cada novo conjunto de pesos é verificada a evolução do erro relativamente ao conjunto anterior. O gradiente do erro indica que aparentemente o treino está no bom caminho, sendo o objectivo encontrar o conjunto de pesos para o qual a derivada da função é zero (ou na prática muito perto de zero). Quando a medida global de erro de treino for menor que um limiar por nós admitido o treino é dado por terminado.

Um outro factor a ter em conta no treino da rede é a taxa de aprendizagem: se para cada passo do treino a variação dos pesos for pequena, a rede será lenta a treinar, por outro lado, taxas altas poderão fazer com que a rede deixe de convergir.

## **3.10 PHP**

O PHP é actualmente umas das linguagens mais difundidas na *web*, para a implementação de *web sites* dinâmicos, para ligação a bases de dados e para a comunicação com outras linguagens (PHP, 2004).

O PHP teve por base o produto PHP/FI (*Personal Home Page Tools/Forms Interpreter*), criada por Rasmus Lerdorf em 1995, sendo inicialmente um conjunto de *scripts* na linguagem Perl para gerir o acesso ao seu *Curriculum Vitae* disponível *on-line*. À medida que novas funcionalidades iam sendo necessárias, Rasmus escreveu uma implementação em C muito mais alargada, a qual tinha funcionalidades de comunicação com bases de dados e possibilitava aos utilizadores o desenvolvimento simples de aplicações *web* dinâmicas. Rasmus decidiu tornar o código do PHP/FI público, para que todas as pessoas o pudessem utilizar e de modo a que outros programadores pudessem corrigir alguns erros e implementar novas funcionalidades (PHP, 2004).

O PHP/FI incluía algumas das funcionalidades básicas do PHP tal como é conhecido hoje. Possuía variáveis semelhantes à linguagem Perl, interpretação automática das variáveis de formulários em HTML (*Hypertext Markup Language*) e podia conter sintaxe HTML.

Em 1997, altura em que surgiu a segunda versão do PHP/FI (2.0), esta linguagem já apresentava um culto de milhares de utilizadores por todo o mundo, estimando-se que cerca de 50.000 domínios tivessem instalado o PHP/FI (sendo esse valor cerca de 1% de todos os domínios existentes nessa altura).

A terceira versão do PHP, 3.0, foi a primeira que se aproximou de uma forma geral do PHP que é conhecido actualmente. Lançado oficialmente em Junho de 1998, foi criado em 1997 por Andi Gutmans e Zeev Suraski, os quais reescreveram completamente o código do PHP/FI 2.0. A sua motivação derivou da necessidade de uma linguagem apropriada para uma aplicação de eCommerce num *web site* a ser desenvolvido para um projecto na Universidade, tendo encontrado no PHP/FI uma base sólida mas não suficiente para o projecto. Em colaboração com Rasmus, Andi e Zeev desenvolveram o que foi oficialmente chamado de PHP 3.0, o sucessor do PHP/FI.

Uma das melhores características do PHP 3.0 é a sua extensa capacidade de expansão. Além de oferecer aos utilizadores finais um sólido conjunto de infra estruturas de ligação a bases de dados, protocolos e APIs (*Application Programming Interfaces*), apresenta possibilidades de expansão de modo a que programadores possam adicionar novos módulos para expandir as suas funcionalidades básicas. Uma das outras inovações introduzidas na versão 3.0 foi o suporte para uma sintaxe orientada a objectos.



Esta nova linguagem foi disponibilizada com um novo nome, retirando a implicação de ser apenas para uso pessoal que o *Personal Home Page* poderia ter. Foi assim simplesmente chamada de PHP, cujo significado está no acrónimo recursivo: PHP: *Hypertext Preprocessor*.

No final de 1998 o número de instalações de PHP subiu para um número estimado de 10% de todos os servidores na internet.

Pouco após a versão 3.0 ter sido lançada, Andi Gutmans e Zeev Suraski começaram a reescrever o núcleo do PHP. Os objectivos passavam pela melhoria do desempenho em aplicações complexas e por melhorar a modularidade da sua base. Foi assim criado um novo núcleo, chamado de *Zend Engine* (derivando dos nomes de Zeev e Andi).

A versão 4.0 do PHP, lançada em Maio de 2000, baseia-se neste novo núcleo e possui um grande número de novas funcionalidades. Além de um muito melhor desempenho global nesta versão, foram introduzidas funcionalidades de suporte para um maior número de servidores *web*, suporte para sessões HTTP e maior segurança na utilização de dados.

O PHP 4.0 é actualmente a última versão, embora já estejam em curso melhoramentos do núcleo *Zend Engine* para a futura versão 5.0.

Actualmente, o PHP é utilizado por centenas de milhares de programadores (valor estimado) e vários milhões de *web sites* possuem-no instalado nos seus servidores, o que corresponderá a cerca de 20% de todos os *sites* na Internet (PHP, 2004).

## 4 Definição do sistema e especificação

Neste capítulo será apresentada a identificação de requisitos do sistema a implementar (protótipo), requisitos esses quer funcionais quer não funcionais. Serão detalhadamente estabelecidas:

- as características e funcionalidades do sistema;
- restrições sobre as quais o sistema será desenvolvido;
- restrições sobre as quais o sistema funcionará.

Em resumo, este capítulo é uma “declaração” escrita das funcionalidades do protótipo e quais as suas características. Não se trata de um capítulo onde será descrita a forma como será feita a implementação dessas mesmas funcionalidades. A descrição detalhada da implementação do sistema e dos seus módulos constituintes será apresentada no Capítulo 5 – Detalhes de implementação e resultados.

### 4.1 Parte I - Visão geral do sistema

A elaboração deste protótipo teve como objectivo o desenvolvimento de um sistema multimédia capaz de captar os sinais fisiológicos de um indivíduo (nomeadamente o batimento cardíaco, a sudação cutânea e a actividade muscular) e utilizar esses sinais para interagir em tempo real com conteúdos multimédia (imagem, vídeo, som e texto). Além disso, e com base nesses mesmos sinais fisiológicos, pretendia-se reconhecer o estado emotivo do indivíduo, reflectindo-se esse estado na interacção com todo o conteúdo do sistema.

O sistema deveria também registar o histórico dos estados emotivos detectados, ou seja, o registo temporal da detecção dos estados emotivos quando o sistema se encontrar em actualização.

#### 4.1.1 Utilização do sistema

O sistema será desenvolvido com vista a ser integrado num *wearable*, um fato construído para o efeito, de modo a que possa ser transportado e utilizado em tempo-real por um indivíduo. O *wearable* possuirá saída visual e sonora através de um monitor TFT (*Thin Film Transistor*) e de colunas colocadas no fato.

O utilizador, através dos seus sinais fisiológicos captados, interagirá com o sistema de uma forma indirecta, não o afectando conscientemente. Uma vez que o sistema se destina a captar os seus sinais fisiológicos e a detectar o seu estado de espírito, permitindo também que esse estado e sinais fisiológicos possam controlar a componente visual e sonora do sistema, o utilizador não terá controlo sobre os seus estados emotivos aquando da utilização do sistema e, consequentemente, controlo directo sobre ele.

O utilizador poderá no entanto, de uma forma directa e consciente, interagir com sistema através da consulta do histórico do sistema (o registo temporal dos estados emotivos detectados).

#### 4.1.2 Arquitectura e interacção com outros sistemas

O sistema será constituído por quatro módulos interligados entre si, cada um desempenhando uma função específica:

- **Módulo de Aquisição dos Sinais Fisiológicos** – Este módulo tem como objectivo a captação dos sinais fisiológicos do indivíduo através de *hardware* (eléctrodos e interface para a captação dos sinais) e software específicos;
- **Módulo de Detecção do Estado Emotivo** – Este módulo é o responsável pela detecção do estado emotivo através de três sinais fisiológicos, sendo esses estados emotivos: a tristeza, o susto, a alegria e o repouso;
- **Módulo de Apresentação** – Este módulo é o responsável pela apresentação visual e sonora do sistema com base no estado emotivo detectado nos três sinais fisiológicos, sendo a apresentação sonora afectada também pelo sinal de vídeo de uma *web cam*;
- **Módulo do Histórico do Sistema** – Este módulo é o responsável pelo registo da data, hora e dos estados emotivos evidenciados pelo utilizador quando a usar o sistema.

Estes módulos constituintes do protótipo a implementar formarão um sistema fechado, não interagindo com qualquer outro software ou sistema.

## **4.2 Parte II – Requisitos Funcionais**

Nesta secção serão descritos, de uma forma mais detalhada e precisa, todas as funcionalidades que o sistema deverá possuir.

### **4.2.1 Descrição das funcionalidades**

Apresentam-se de seguida os requisitos funcionais divididos por diferentes áreas, áreas estas referentes aos módulos constituintes do sistema, ao *wearable* e às ferramentas, linguagens e ambientes de programação

#### **Módulo de Aquisição dos Sinais Fisiológicos**

- deverá permitir a captação do sinal relativo ao batimento cardíaco (HR), actividade muscular (EMG) e condutividade eléctrica da pele (GSR). Estes sinais, como referido no Capítulo 2 – Emoções e Computação Afectiva, são suficientes para a detecção do estado emotivo de um indivíduo;
- deverá permitir a transmissão em tempo real desses mesmos sinais para outros sistemas através de um protocolo normalizado – estes sinais necessitam de ser transmitidos para outro sistema de forma a que se possa proceder à sua análise e à detecção do estado emotivo;
- o sistema deverá permitir a exportação dos dados captados e gravados em formatos normalizados (como ficheiros de texto \*.txt), de modo a que possam ser utilizados por outros módulos ou sistemas.

#### **Módulo de Detecção do Estado Emotivo**

- o sistema deverá detectar os seguintes estados emotivos: repouso, susto, tristeza e alegria;
- a análise para detecção do estado emotivo deverá ser feita com um intervalo de 20 segundos, tendo por base os dados fisiológicos captados em tempo-real;

- o sistema não detectará, dentro de um estado emotivo, diferentes níveis para esse estado – ou seja, o sistema apenas detectará o estado de alegria e não diferentes graus de alegria;

### **Módulo de Apresentação**

- este módulo deverá permitir a entrada em tempo-real de, através de um protocolo normalizado, de 3 sinais relativos a dados fisiológicos: batimento cardíaco (HR), actividade muscular (EMG) e condutividade eléctrica da pele (GSR);
- este módulo deverá ser constituído por sub-módulos (ou Pacotes), cada um relativo a um dos estados emotivos que se pretende detectar: alegria, tristeza, susto e repouso – os estados básicos e mais gerais das emoções humanas;
- o Pacote relativo ao estado da tristeza será constituído por 3 Pacotes, de modo a representar diferentes graus dessa emoção;
- além dos Pacotes relativos a cada uma das emoções, o sistema deverá possuir um Pacote de entrada no sistema;
- os Pacotes (exceptuando o Pacote relativo à entrada no sistema) deverão permitir a manipulação em tempo real da componente visual e/ou sonora através da entrada dos sinais fisiológicos e/ou de entrada do sinal de vídeo;
- quando for detectado determinado estado emotivo deverá ser mostrado pelo sistema o Pacote relativo a esse estado, durante o tempo em que esse estado emotivo se mantenha, sendo o tempo mínimo de permanência em cada Pacote de 1 minuto;
- cada um dos Pacotes (exceptuando o Pacote relativo à entrada no sistema) deverá ter um conjunto com no mínimo 2 excertos de poemas alusivos a determinada emoção, excertos esses que serão mostrados aleatoriamente quando o Protótipo esteja a ser apresentado;
- cada Pacote (exceptuando o Pacote relativo à entrada no sistema) deverá possuir uma caixa de texto com o nome do sistema, a data e a hora, um excerto de um poema e o sinal de vídeo em tempo-real captado através de uma *web cam*;

- cada um dos Pacotes deverá ter imagens, vídeos, texto e som diferentes para cada um deles, estando esses conteúdos relacionados com o estado emotivo a ser apresentado pelo Pacote;
- o sinal de vídeo da *web cam* deverá ser utilizado em pelo menos um Pacote para a criação de música interactiva, ou seja, a produção de música com base no sinal de vídeo captado em tempo-real;
- os sinais fisiológicos serão analisados com o objectivo de detectar o estado emotivo e deverão em tempo real afectar a apresentação visual e sonora do sistema;
- deverá ser possível a utilização de vários modos de produção sonora, quer através de síntese de som própria, da utilização de *plug-ins* de instrumentos/sintetizadores virtuais MIDI ou da utilização de fontes sonoras externas ao sistema.

#### **Módulo do Histórico do Sistema**

- o sistema deverá permitir a gravação de um registo temporal dos estados emotivos detectados, para uma posterior visualização ou utilização – a cada alteração detectada do estado emotivo, o sistema deverá registar essa alteração, e a data e a hora em que ocorreu, para que esse mesmo registo possa mais tarde ser consultado ou submetido a uma possível análise estatística.

#### ***Wearable***

- o sistema a implementar deverá ser montado e demonstrado num fato desenhado para o efeito (um *wearable*);
- o *wearable* deverá possuir um ecrã onde será apresentado o Pacote relativo ao estado emotivo do indivíduo (do Módulo de Apresentação), colunas para o som e os eléctrodos necessários para a captação dos sinais fisiológicos.

#### **Outros Requisitos**

- o sistema deverá ser implementado na plataforma Windows Millenium Edition (Me) ou superior – a plataforma a que mais

facilmente se tem acesso, a mais difundida e a que oferece uma maior base comum para o funcionamento de várias ferramentas de programação, edição de multimédia (áudio, vídeo e imagem) e sistemas de aquisição de sinais fisiológicos;

- o Módulo de Apresentação deverá ser implementado numa linguagem de programação visual (Pure Data ou Max/MSP) – estas linguagens de programação são orientadas e optimizadas para a produção visual e sonora, possuindo um elevado número de objectos para a manipulação de vídeo, áudio, imagens e MIDI em tempo real, podendo todos estes elementos ser integrados num único sistema. Uma vez que a plataforma de implementação é Windows XP e Me, será utilizada a linguagem Pure Data.
- O Módulo de Aquisição de Sinal deverá ser implementado através da utilização do *hardware* WaveRider Jr. e do software que o acompanha, o WaveWare, juntamente com os eléctrodos fornecidos.
- Os restantes Módulos poderão ser implementados em qualquer linguagem/ferramenta disponível para a plataforma Windows.

#### **4.2.2 Resposta do sistema**

O sistema terá de responder em tempo real à alteração aos diversos dados de entrada (*web cam* e sinais fisiológicos captados) uma vez que há Pacotes no Módulo de Apresentação nos quais a alteração da componente visual e sonora deverá depender directamente desses sinais de entrada.

#### **4.2.3 Calendarização da implementação do sistema**

A implementação do sistema será iniciada com o Módulo de Apresentação, seguindo-se o Módulo de Aquisição dos Sinais Fisiológicos, o Módulo de Detecção do Estado Emotivo e finalmente o Módulo do Histórico do Sistema.

Esta ordem de implementação deve-se ao facto de os Módulos de Detecção do Estado Emotivo e Histórico do Sistema dependerem da implementação prévia dos Módulos de Apresentação e Aquisição

dos Sinais Fisiológicos. O Módulo de Aquisição dos Sinais Fisiológicos depende por sua vez da disponibilização prévia do WaveRider e do WaveWare para a sua implementação. Não estando o Módulo de Apresentação dependente de nenhum outro módulo ou *hardware/software* específico, será este o primeiro a ser implementado.

O projecto e a construção do *wearable* não depende da prévia implementação de qualquer Módulo do sistema, daí que a sua implementação possa decorrer em qualquer fase do desenvolvimento do protótipo, devendo estar, no entanto, concluído no final da implementação do último módulo, Módulo do Histórico do Sistema, para que o sistema possa ser utilizado e testado em tempo-real utilizando o *wearable*.

### 4.3 Arquitectura do sistema

Como se referiu no Capítulo 4.1.3., o sistema a implementar será constituído por 4 módulos, cada um desempenhando uma função específica e interligados entre si através de vários protocolos normalizados:

- **Módulo de Aquisição dos Sinais Fisiológicos** – Este módulo, após a captação dos sinais fisiológicos do indivíduo (batimento cardíaco, condutividade eléctrica da pele e actividade muscular), deverá convertê-los para que possam ser transmitidos através de um protocolo normalizado para outros sistemas/módulos. Todo este processo deverá ser feito em tempo-real;
- **Módulo de Apresentação** – Este módulo tem como entrada os 3 sinais fisiológicos recolhidos e convertidos pelo Módulo de Aquisição dos Sinais Fisiológicos. Possui também como entrada o sinal de vídeo de uma *web cam*. Após a recepção dos sinais fisiológicos, este Módulo transmite-os ao Módulo de Detecção do Estado Emotivo a fim de ser detectado o estado emotivo do utilizador. Com base nesse estado o sistema apresentará a informação sonora e visual de acordo com esse mesmo estado, além de continuamente essa mesma informação sonora e visual poder ser alterada através dos sinais fisiológicos e de vídeo;



- **Módulo de Detecção do Estado Emotivo** – Este módulo terá como entrada os 3 sinais transmitidos pelo Módulo de Apresentação. Neste módulo será feito um processamento sobre esses dados de modo a detectar o estado emotivo do indivíduo. Após a detecção desse estado, o resultado será transmitido ao Módulo de Apresentação;
- **Módulo do Histórico do Sistema** – Este módulo regista o novo estado emotivo detectado, juntamente com a data e hora em que ocorreu.

Apresenta-se na Figura 4.1 uma representação da interligação dos 4 Módulos.

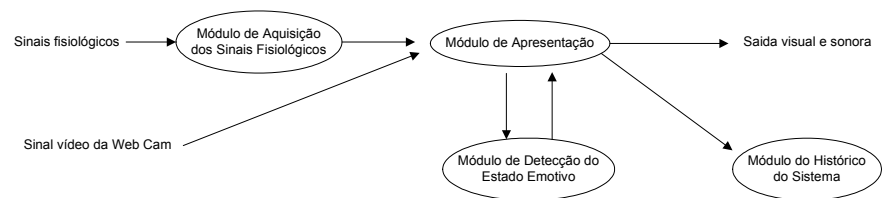


Figura 4.1 – Interligação dos Módulos constituintes do sistema

## 5 Detalhes de implementação e resultados

Neste capítulo serão descritas em pormenor as várias fases necessárias à implementação do protótipo. Descrever-se-ão os detalhes de implementação dos módulos constituintes e a ligação entre eles. Apresentar-se-á também a aplicação do protótipo num *wearable*, um fato desenhado e construído especificamente para o efeito, assim como os resultados observados aquando da utilização do sistema a ser utilizado em tempo real.

### 5.1 Arquitectura do Sistema

O sistema é constituído por vários módulos, cada um desempenhando uma função específica e recorrendo a tecnologias distintas:

- **Módulo de Apresentação** – módulo desenvolvido em PD/GEM, sendo constituído por diversos módulos mais pequenos (ou Pacotes), responsáveis pela apresentação gráfica e sonora dos estados emotivos do utilizador;
- **Módulo de Aquisição dos Sinais Fisiológicos** – este módulo é constituído pelo *hardware* WaveRider Jr. e pelo software WaveWare, sendo responsável pela aquisição do sinal muscular, do batimento cardíaco e da condutividade eléctrica da pele;
- **Módulo de Detecção do Estado Emotivo** – módulo desenvolvido em Java e Matlab, no qual uma rede neuronal detecta o estado emotivo do utilizador com base nos sinais fisiológicos captados pelo WaveWare;
- **Módulo do Histórico do Sistema** – módulo desenvolvido em Java, MySQL e PHP e responsável pelo registo numa base de dados da data, hora e dos estados emotivos evidenciados pelo utilizador.

Todos os módulos encontram-se interligados entre si através de diversos protocolos:

- O Módulo de Aquisição dos Sinais Fisiológicos e o Módulo de Apresentação estão interligados através de MIDI;

- O Módulo de Apresentação e o Módulo de Detecção do Estado Emotivo comunicam através de *sockets* em Java;
- O Módulo de Apresentação e o Módulo do Histórico do Sistema comunicam através de *sockets* em Java.

Na Figura 5.1 mostra-se uma representação gráfica detalhada da interligação entre os Módulos do sistema.

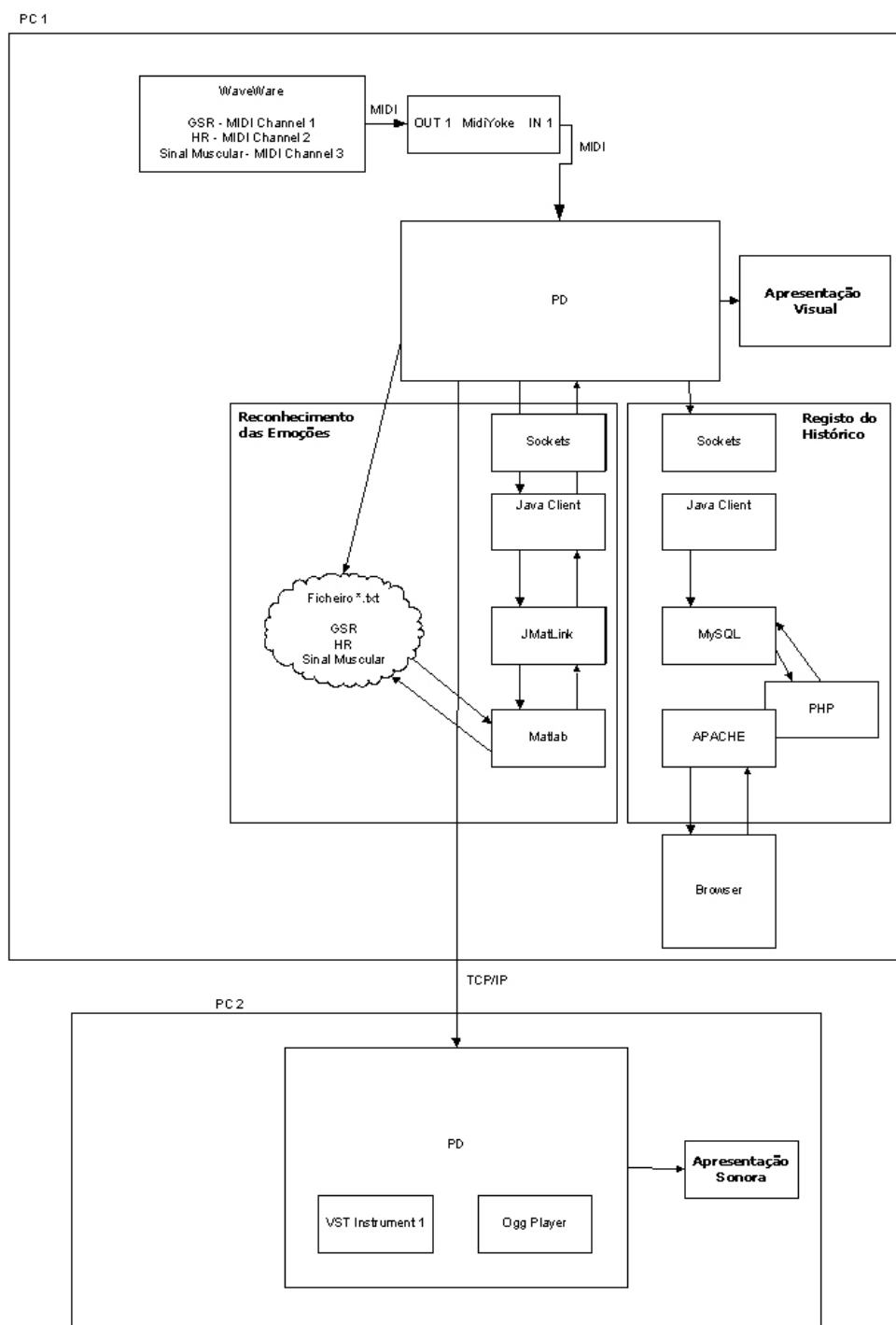


Figura 5.1 – Detalhe da interligação dos Módulos constituintes do sistema

A descrição das funcionalidades e os detalhes inerentes a cada um dos módulos, assim como os detalhes da interligação entre eles são apresentados nas próximas secções.

## 5.2 Módulo de Apresentação

O desenvolvimento do protótipo teve início com a implementação do Módulo de Apresentação, mais concretamente, com o desenvolvimento de cada um dos sub-módulos que o constituem, representando cada um deles uma das emoções que se pretendiam detectar. A estes sub-módulos, e por uma questão de apresentação e de legibilidade, convencionou-se chamar de “Pacotes”, termo que será utilizado doravante. Assim, estes Pacotes não são mais do que pequenos módulos que no seu conjunto formam o Módulo de Apresentação. Este módulo é então constituído pelos seguintes Pacotes:

- Pacote de Entrada;
- Pacote da Tristeza;
- Pacote do Susto;
- Pacote da Alegria;
- Pacote do Repouso.

Todos os Pacotes foram implementados em PD/GEM com uma resolução de 800x600, tendo alguns deles como característica a interactividade em termos visuais e sonoros com base nos sinais fisiológicos enviados pelo software de aquisição de dados, o WaveWare (para além de serem activados consoante a análise e a detecção dos estados emotivos com base nesses mesmos sinais).

Estes Pacotes, exceptuando o Pacote de Entrada (que será apresentado mais à frente), possuem em comum o facto de apresentarem no canto inferior esquerdo uma caixa contendo a seguinte informação:

- o nome do protótipo (*from within*);
- a hora e a data;
- excertos de poemas alusivos ao tema de cada Pacote (tristeza, susto, alegria e repouso);
- a entrada em tempo real do sinal de vídeo de uma *web cam*.

Cada um desses Pacotes possui um número variado de excertos de poemas que são apresentados aleatoriamente quando o Pacote é carregado.

Como exemplo, apresenta-se na Figura 5.2 a caixa relativa a um poema aleatório no Pacote do Repouso.

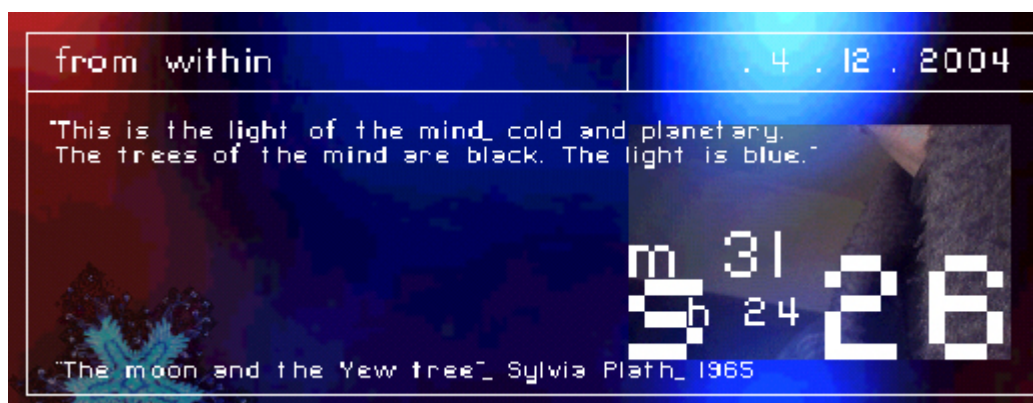


Figura 5.2 – Caixa de informação com excerto de um poema no Pacote do Repouso

De seguida apresenta-se com algum detalhe o funcionamento de todos os Pacotes, assim como a interligação entre eles e o WaveWare.

## 5.2.1 Pacotes

### Pacote de Entrada

Este Pacote é apenas apresentado quando se dá início ao funcionamento do protótipo, tendo como finalidade servir como “porta de entrada” para a apresentação das emoções. A partir do momento em que é detectada uma emoção, o Pacote de Entrada deixa de ser visível dando lugar ao Pacote correspondente a essa emoção.

O Pacote é constituído por uma série de imagens de portas a serem substituídas continuamente e em ciclo por outras portas da mesma dimensão, apresentando cada uma das portas uma velocidade de substituição diferente. Cada porta pretende representar a entrada para uma determinada emoção (embora neste protótipo exista um número superior de portas em relação ao número de emoções detectadas).



Figura 5.3 – Conjunto de portas presentes no Pacote de Entrada

Quando uma emoção é detectada, todas as portas são substituídas por cortinas de veludo vermelho durante 3 segundos. Após esse tempo, desaparecem todas as cortinas, mantendo-se apenas a cortina sobre a porta correspondente à emoção que foi detectada durante mais 3 segundos. Findo este tempo é apresentado o Pacote da emoção detectada, desaparecendo a única cortina restante (Figura 5.4).



Figura 5.4 – Sequência de cortinas mostradas na transição para um novo estado emotivo

A transição para cada um dos outros Pacotes de emoções é semelhante à passagem do Pacote de Entrada para um dos Pacotes que representa uma emoção. Ou seja, a transição entre todos os Pacotes é feita com a apresentação das portas cobertas com cortinas de veludo durante 3 segundos, seguindo-se o desaparecimento das mesmas excepto a correspondente à emoção

que vai ser apresentada, mantendo-se esta única cortina durante mais 3 segundos até ser apresentado um novo Pacote.

### **Pacote da Tristeza**

O Pacotes da Tristeza é constituído por 3 Cenas, pretendendo cada uma delas representar um diferente nível de tristeza que o indivíduo possa evidenciar.

Embora se tenha dividido o Pacote da Tristeza em 3 níveis, na realidade, a rede neuronal não detecta diferentes graus de determinada emoção, neste caso, diferentes graus de tristeza. O que se pretendeu exemplificar neste pacote é que quando determinado indivíduo se mantém neste estado durante um certo período de tempo, o grau dessa emoção poderá ser mais pronunciado à medida que o tempo avança. Dividiu-se assim este Pacote em 3 Níveis, pretendendo cada um deles representar um diferente grau de tristeza.

O tempo que se convencionou neste protótipo, para efeitos de demonstração, foi de um mínimo de 60 segundos de permanência em cada Nível antes de passar ao Nível superior (Nível 1 para o Nível 2, e do Nível 2 para o Nível 3). Assim, se for apresentado o Nível 1 e o indivíduo permanecer no estado de tristeza durante mais de 60 segundos (ou seja, duas vezes consecutivas de detecção desse estado, uma vez que a frequência de detecção é feita de 20 em 20 segundos), é apresentado o Nível 2. O processo é o mesmo para a passagem do Nível 2 para o 3.

O Nível 1 é sempre a primeira Cena a ser apresentada quando se detecta um estado de tristeza se o estado imediatamente anterior for de uma outra qualquer emoção. O Nível 1 é assim a “porta de entrada” para o estado da tristeza. Caso o indivíduo permaneça durante 60 segundos neste estado é apresentado o Nível 2. Se o indivíduo continuar mais 60 segundos no Nível 2 apresenta-se por fim o Nível 3, mantendo-se nesta Cena enquanto o estado detectado for o da tristeza. Em qualquer destes Níveis, caso seja detectado um outro estado que não o da tristeza, é apresentado o Pacote correspondente a esse novo estado detectado e repetindo-se todo este processo quando for detectado um novo estado de tristeza.

- *Pacote da Tristeza Nível 1*

Em termos visuais o Pacote é constituído por uma série de imagens de nuvens brancas num céu azul, aparecendo e desaparecendo cada uma delas em gradação de visibilidade (*fade-in* e *fade-out*), sendo este efeito conseguido no GEM através da alteração dos níveis de *alpha* de cada uma dessas imagens. Assim, todas as imagens estão presentes em qualquer altura, só que enquanto algumas se encontram visíveis, outras estão invisíveis, uma vez que o nível de *alpha* vai sendo alterado independentemente em cada uma das imagens.

Este módulo possui um texto central, uma citação de Paul Bowles do filme “*The Sheltering Sky*” de Bernardo Bertollucci, que vai correndo em *scroll* no centro do Pacote, da direita para a esquerda. Este texto aparece durante todo o tempo em que o Pacote esteja activo, repetindo-se o *scroll* quando a citação chega ao fim.

O som apresentado neste Pacote foi retirado da banda sonora do mesmo filme, “*The Sheltering Sky*”. Trata-se de um *loop* convertido para o formato \*.ogg, repetindo-se enquanto o Pacote está activo. Optou-se pela utilização do formato \*.ogg em detrimento do \*.wav uma vez que o \*.ogg é um formato comprimido, semelhante ao \*.mp3, sendo portanto mais leve em termos de utilização de recursos que um ficheiro de som não comprimido \*.wav. Optou-se também pelo \*.ogg, em detrimento do \*.mp3 uma vez que o objecto disponibilizado pelo PD para a leitura de \*.ogg’s é bastante mais flexível que o de leitura de \*.mp3’s. O *patch* responsável pela reprodução de *loops* apresenta-se na Figura 5.5.

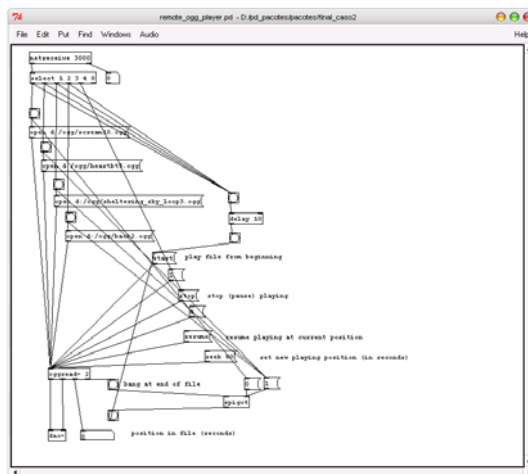


Figura 5.5 – *Patch* remote\_ogg\_player.pd





Figura 5.6 – Nível 1 do Pacote da Tristeza

#### - Pacote da Tristeza Nível 2

O segundo nível da tristeza representado neste Pacote pretende retratar um estado de tristeza mais forte que o primeiro, encontrando-se o indivíduo num estado já mais melancólico.

Em termos visuais, o Pacote é constituído por um vídeo montado com base em imagens de um prédio em demolição, rodeado por escombros e entulho, sobre um fundo negro. Existem também um conjunto de imagens de 5 Santas que aparecem e desaparecem sobre o vídeo através da variação do nível de alfa das imagens, tal como acontece no Pacote do primeiro nível da tristeza.

No que respeita ao som, o Pacote apresenta um loop de uma composição de Bach, convertida para o formato \*.ogg.



Figura 5.7 – Nível 2 do Pacote da Tristeza

### - Pacote da Tristeza Nível 3

No terceiro nível da tristeza pretende-se representar um estado ainda mais profundo desse sentimento, quando um indivíduo se encontra num estado em que as lágrimas poderão cair, tal como a representação visual o sugere.

Assim, em termos visuais, este Pacote apresenta uma montagem feita com base numa fotografia de Man Ray de um olho feminino com uma lágrima a cair, sobre um fundo de cetim vermelho vivo. Sobre esta composição visual, e com a imagem de base da lágrima, existem duas imagens de lágrimas que se “desprendem” da fotografia original e que caem repetidamente.

Este Pacote não possui som.

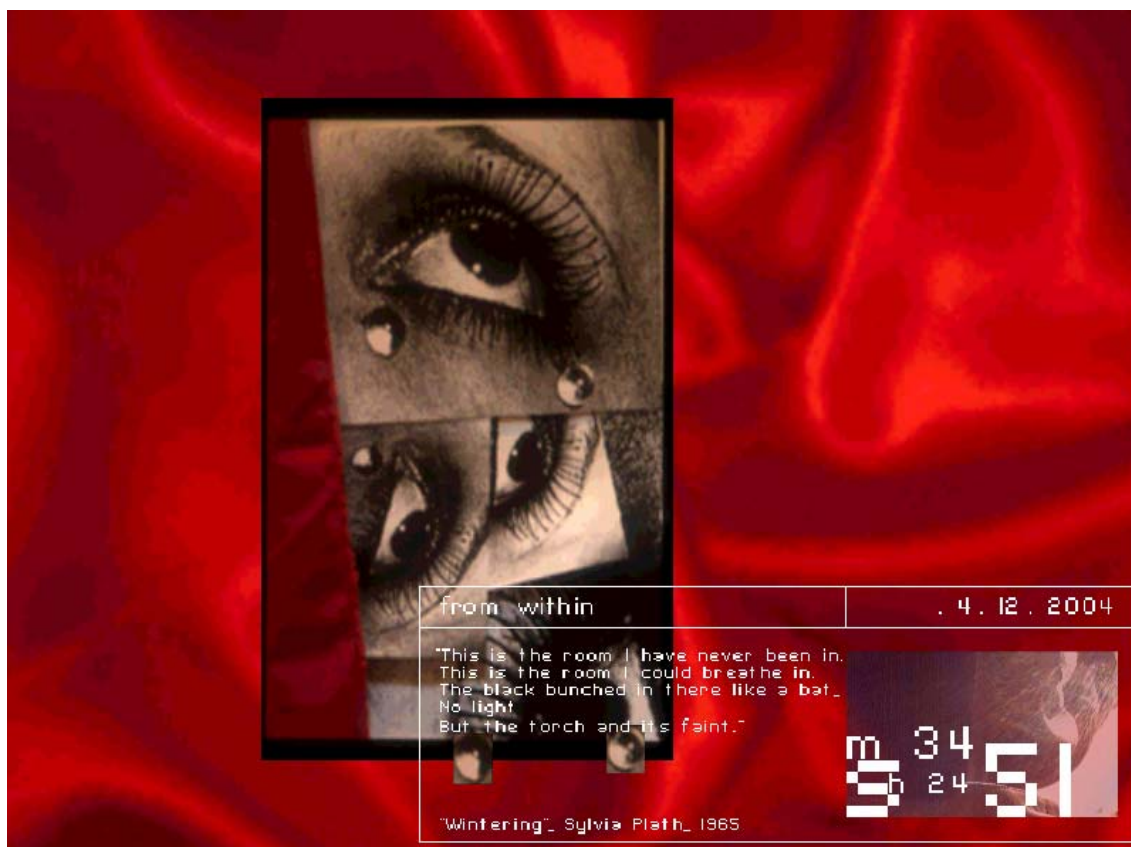


Figura 5.8 – Nível 3 do Pacote da Tristeza

### **Pacote do Susto**

Este Pacote é activado quando se detecta um estado de susto ou de excitação súbita do indivíduo. Em termos visuais, o Pacote é constituído por uma colagem de várias imagens de pessoas, dispostas em 3 tiras horizontais que se movimentam em scroll alternado para cada tira (a tira do topo possui um *scroll* da direita para esquerda, a central da esquerda para a direita e a de baixo da direita para a esquerda). Cada uma dessas 3 tiras de pessoas vai alternando por sua vez para uma outra tira com as mesmas pessoas mas com a cara oculta por imagens de várias figuras de desenhos animados e super-heróis.

Sobre estas tiras, estão ainda presentes 3 vídeos de dimensão 320x240 dispostos aleatoriamente quando o Pacote é apresentado pela primeira vez. Um destes vídeos é a imagem em tempo real captado pela web cam. Estes três vídeos tornam gradualmente visíveis e invisíveis (*fade-in* e *fade-out*) através da variação dos seus níveis de alfa. À terceira vez que cada um dos vídeos se torna

invisível, a sua posição no ecrã muda aleatoriamente, repetindo-se o processo de *fade-in* e *fade-out* durante três vezes até mudarem novamente de posição e assim sucessivamente.

A velocidade da alternância das imagens das tiras, assim como a posição dos 3 vídeos e a frequência da variação dos seus níveis de alfa é controlada pelo batimento cardíaco, sendo esta variação tanto mais rápida quanto mais elevado for o batimento.

A componente sonora divide-se em duas partes: na fase inicial, quando o Pacote é preparado para ser apresentado, e ainda na fase de transição em que aparecem as cortinas de veludo, é tocado um pequeno som de uma porta a ranger seguido de um grito. Quando se entra finalmente no Pacote do Susto, o som que se apresenta é o de uma gravação do batimento cardíaco convertida para o formato \*.ogg.

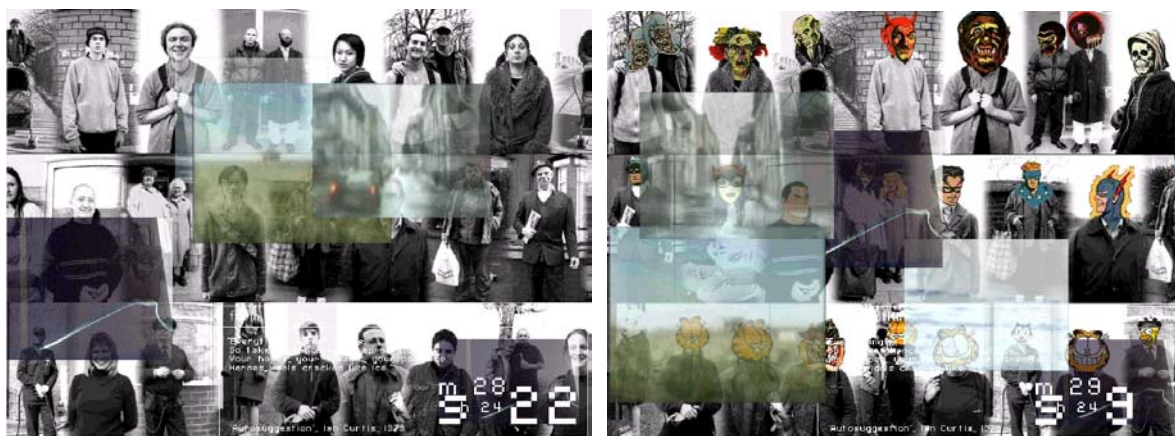


Figura 5.9 – Alternância das tiras de imagens constituintes do Pacote do Susto

### **Pacote da Alegria**

Este Pacote, apesar de se tratar apenas de um único Pacote isolado, é constituído por 2 Cenas.

Na primeira cena, é apresentado um vídeo de flores ao qual é aplicado um efeito de mosaico que o divide em 4 rectângulos horizontais por 4 verticais. Sobre esse vídeo é apresentada uma sequência de 4 imagens de flores a branco que vão alternando entre si de acordo com a variação dos seus níveis de alfa. Assim, a após a apresentação da primeira flor sobre o vídeo, o seu nível de alfa decresce desde o valor máximo até ao valor mínimo, dando lugar à

segunda flor cujo nível de alfa aumenta desde o mínimo até ao máximo, e assim sucessivamente.

Quando se verifica o aumento do batimento cardíaco acima das 80 pulsações por minuto, o Pacote apresenta a segunda Cena. Caso à entrada do Pacote da Alegria corresponda já a um batimento cardíaco acima das 80 pulsações, o módulo entra sempre na primeira cena, passando de seguida para a segunda mas permanecendo sempre na primeira cena um tempo mínimo de 10 segundos.

A segunda Cena apresenta uma série de vídeos e imagens de flores dispostos de forma geométrica. Sobre esses vídeos e imagens de flores apresentam-se outras imagens de flores que vão aumentando de tamanho, simulando o seu crescimento. Quando cada flor atinge o seu tamanho máximo, desaparece voltando ao seu tamanho mínimo e voltando a repetir-se o ciclo de crescimento, mas numa nova posição no ecrã que é definida aleatoriamente. A velocidade de crescimento e o tamanho máximo das flores são afectados pelo batimento cardíaco, sendo a velocidade de crescimento tanto maior quanto mais rápido for o batimento detectado.

Este módulo possui som produzido através do objecto [vst~], objecto esse que permite a inclusão de um sintetizador virtual VST (*Virtual Studio Technology*) num *patch* PD e o controlo dos seus recursos como se de outro normal objecto PD se tratasse.

O sintetizador VST escolhido para o protótipo foi o Native Instruments FM7, sendo utilizado neste Pacote o *patch Electric Harp* (FM7, 2004). O som do sintetizador é produzido de acordo com o valor batimento cardíaco. Estando os valores do batimento cardíaco compreendidos, normalmente, entre as 30 e as 150 pulsações por minuto, fez-se um mapeamento em PD desta gama de valores para a gama 73 até 97, representando esta gama as notas MIDI do sintetizador. Assim, o valor 73 corresponde à oitava imediatamente acima do Dó central e o valor 97 a 3 oitavas acima do Dó central. A escolha desta gama deveu-se ao próprio som do *patch* escolhido: com valores inferiores a 73 produz-se som muito grave, acima de 97 sons demasiado agudos. A gama escolhida produz assim os sons mais adequados para o efeito pretendido.





Figura 5.10 – FM7, Sintetizador Virtual VST da Native Instruments (FM7, 2004)

Os valores à saída da nova gama são ainda tratados pela abstracção PD “Scale”, antes de serem enviados ao sintetizador virtual FM7. Esta abstracção serve como um “ímã” que desvia os valores originais da entrada para os valores mais próximos de uma escala musical, tendo-se neste caso, e por se tratar do Pacote da Alegria, escolhido a escala de Dó Maior. Assim, caso o valor de entrada corresponda a uma nota de uma escala menor, ou qualquer outra escala, é substituído pelo valor mais próximo que pertença à escala de Dó Maior. A saída da abstracção Scale é então enviada para o FM7.

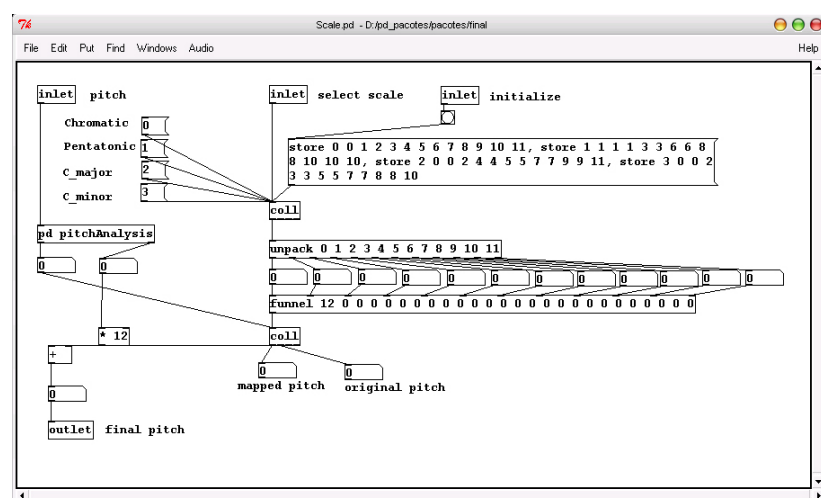


Figura 5.11 – Patch Scale.pd responsável pela criação de uma melodia MIDI em determinada escala

A frequência de envio dos valores de saída da abstracção Scale para o FM7 é condicionada também pelo batimento cardíaco. À medida que o batimento cardíaco aumenta, os valores são enviados com uma maior frequência para o FM7, tornando assim a música mais rápida. Pelo contrário, se o valor do batimento cardíaco diminui, a frequência do envio das notas diminui igualmente, tornando a música mais lenta.

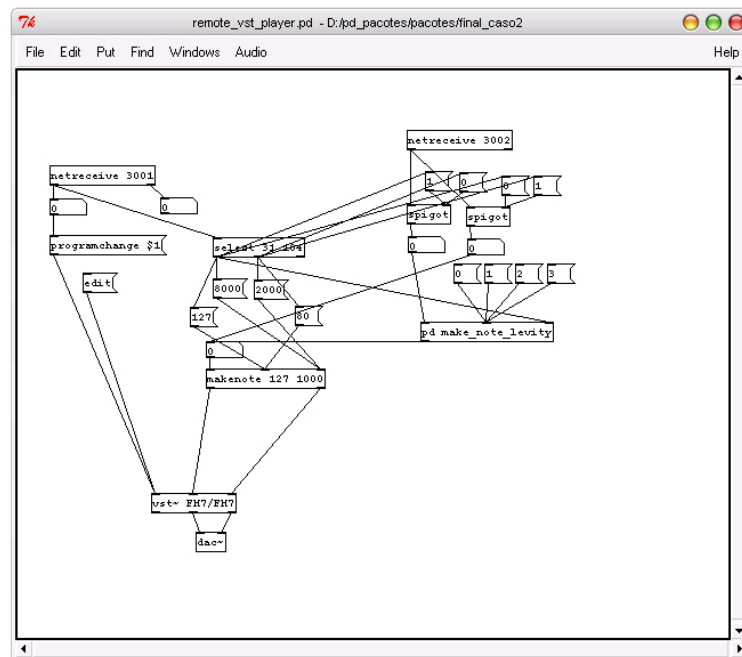


Figura 5.12 – Patch remote\_vst\_player.pd para a inclusão de instrumentos VST no PD

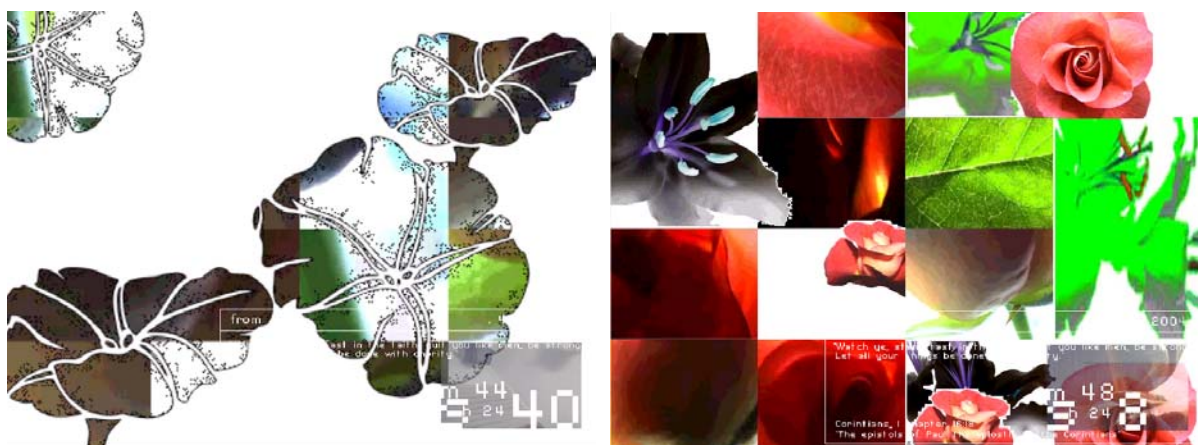


Figura 5.13 – Sequência de entrada no Pacote da Alegria

### **Pacote do Repouso**

Este Pacote é apresentado quando se detecta que o indivíduo se encontra num estado de repouso ou de calma. Em termos visuais, o Pacote é marcado fortemente por um ambiente etéreo, abstracto.

Possui um vídeo com fortes componentes azul, vermelho e preto, lembrando bolhas e formas arredondadas abstractas. Em termos de imagens, possui uma série de elementos gráficos remiiscentes de pequenas bactérias e de formas simétricas e perfeitas. Estes elementos gráficos movimentam-se verticalmente no ecrã, sendo a sua velocidade de deslocamento afectada pelo valor do batimento cardíaco: quando mais rápido este valor, mais rapidamente estes elementos gráficos se movimentam.

Este módulo possui também som produzido através de um sintetizador virtual VST FM7 e do *patch* \* *Ayusca* \* (FM7, 2004). Esse som é produzido através da análise do histograma da componente de azul da entrada da *web cam*. A escolha da componente de azul deve-se à sua predominância sobre as outras componentes quando da utilização da *web cam* à luz do dia e em espaços abertos.

O histograma da *web cam* relativo à componente de azul possui as intensidades entre 0.000 e 1.000 ao longo do nível de azul (de 0, o mais escuro, a 255, o mais claro). Após a análise visual do histograma do vídeo, verificou-se que o nível predominante de azul se situa na zona do valor 50, razão pela qual se escolheu este valor e a sua intensidade como fonte de produção musical.

A escala das intensidades, compreendida entre 0.000 e 1.000, é assim convertida, tal como no Pacote da Alegria, para uma nova escala, neste caso de 49 a 61, correspondendo o 49 ao valor da nota Dó na oitava imediatamente anterior ao Dó central (61). A escolha desta nova escala deve-se, tal como no Pacote da Alegria, às próprias características do *patch* escolhido do FM7, sendo esta a gama de notas que produz os melhores resultados em termos sonoros.



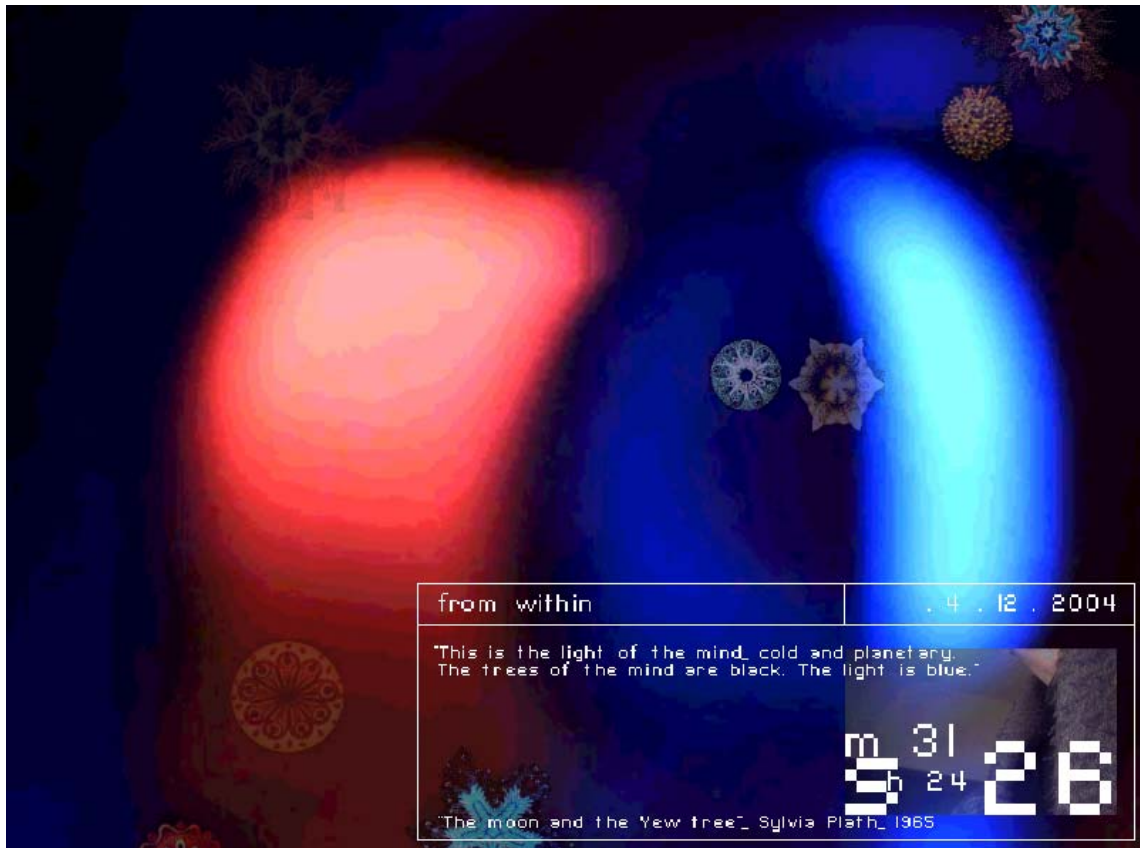


Figura 5.14 – Pacote do Repouso

### 5.2.2 Distribuição em rede do sistema

À medida que a implementação em PD do Módulo de Apresentação e dos seus Pacotes constituintes ia evoluindo, e com a crescente inclusão de imagens e vídeos, verificou-se que o protótipo ficava bastante lento quando se incluiu a componente de áudio (o sintetizador virtual para a criação sonora interactiva e objectos para reprodução de amostras de som).

Como a reprodução de amostras de som e a criação de música interactiva através do sintetizador virtual são componentes fundamentais do Módulo de Apresentação, foi necessário distribuir o sistema por outro computador, para que fosse possível obter um melhor aproveitamento de recursos.

Os *patches* do protótipo estão assim distribuídos por dois computadores: um dedicado ao processamento de áudio e outro dedicado a toda a componente gráfica, aquisição dos sinais fisiológicos do WaveWare, detecção do estado emotivo em Matlab e registo do histórico através de Java e PHP e MySQL.

A comunicação entre os dois computadores em rede é feita através de TCP, utilizando um simples cabo cruzado e os objectos [netsend] e [netreceive] disponibilizado no PD para o efeito, permitindo assim que dois computadores ligados em rede com patches em PD possam comunicar entre si. A ligação iniciada pelo computador que quer enviar mensagens para outro através de [netsend] tem de indicar o endereço IP do outro computador e um número do porto para onde as mensagens serão enviadas. No lado do computador que as recebe através de [netreceive] tem que se especificar o mesmo número do porto para que se possa receber as mensagens enviadas.

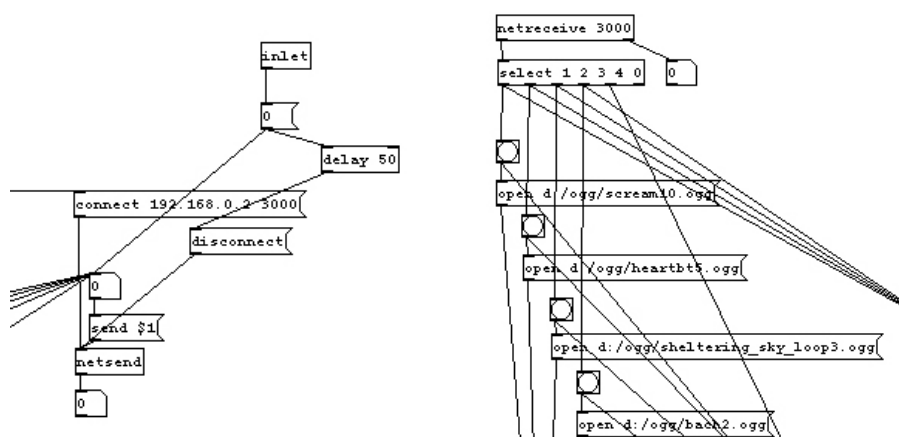


Figura 5.15 – Utilização de [netsend] e [netreceive] para a comunicação entre patches distribuídos por 2 computadores

A comunicação no protótipo é unidireccional, sendo o computador dedicado ao áudio meramente um reproduzidor de som. Este computador recebe informação do computador com o sistema principal sobre que excerto de áudio deverá reproduzir, por quanto tempo o deverá reproduzir, quando deverá parar e, no caso do sintetizador virtual, qual a nota que deverá tocar, com que som do sintetizador e durante quanto tempo. Não existe comunicação do computador reproduzidor de áudio para o computador de controlo.

### 5.3 Módulo de Aquisição dos Sinais Fisiológicos

Este Módulo é o responsável pela aquisição dos sinais fisiológicos necessários para a interacção com o sistema em tempo real e para o reconhecimento das emoções.

### 5.3.1 WaveRider Jr. e WaveWare

O WaveWare, o software fornecido com o WaveRider Jr., foi utilizado para a detecção e captação em simultâneo de 3 sinais: GSR, HR e o sinal muscular.

O *hardware* do WaveRider Jr. possui 2 canais gerais para a captação de sinais (o canal A e B) e um outro canal dedicado à captação do sinal GSR (o canal GSR). Cada um destes sinais é captado através de eléctrodos os quais se colocam no corpo em configurações e locais específicos, de modo a que cada um desses sinais possa ser obtido da forma mais fiável e tentando minimizar a presença de ruído. Para a colocação correcta de cada um dos eléctrodos seguiram-se as indicações apresentadas no manual do WaveRider, do qual foram retiradas as figuras apresentadas de seguida (MindPeak, 2003b).

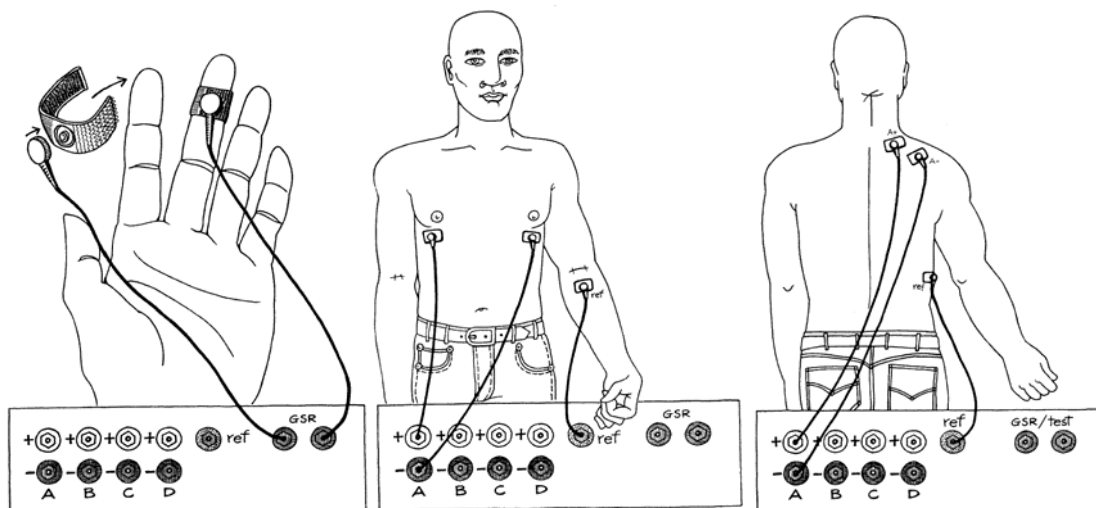


Figura 5.16 – Colocação correcta dos eléctrodos para o sinal GSR, HR e muscular (MindPeak, 2003b)

Além da colocação correcta dos eléctrodos, foi necessário configurar o WaveWare para que captasse correctamente os sinais de entrada. Uma vez que cada uma das entradas gerais (A e B) pode ser configurada para qualquer sinal (HR, muscular e ondas cerebrais), teve que se configurar correctamente cada uma dessas entradas para os sinais HR e muscular. No caso do GSR, uma vez que a entrada é dedicada a esse sinal, apenas foi necessário configurar a gama de valores de entrada que se pretendia captar (*Range of GSR*).

A Figura 5.17 apresenta as caixas de diálogo do WaveWare nas quais se faz a configuração dos parâmetros para cada um dos sinais captados (GSR, HR e muscular):

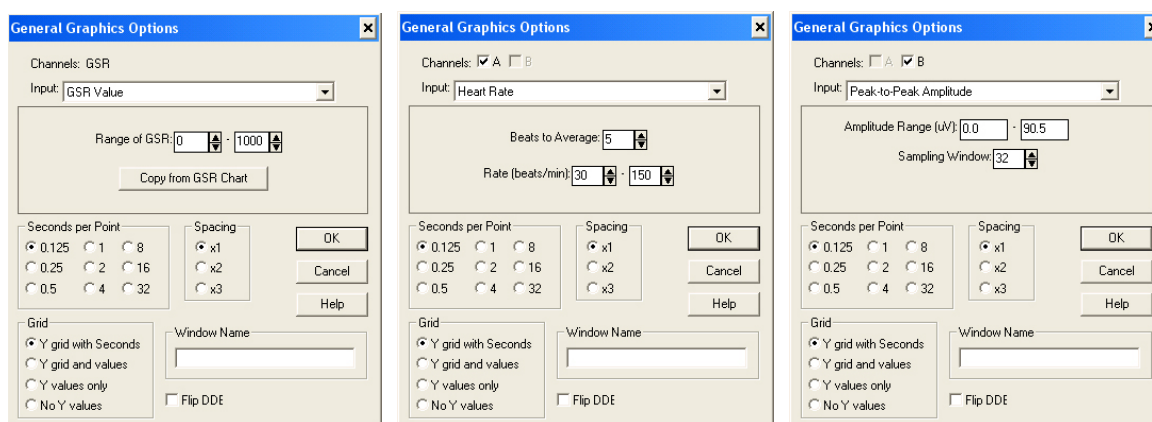


Figura 5.17 – Opções do WaveWare para a configuração do sinal GSR, HR e muscular

Os parâmetros para cada um dos sinais foram configurados de acordo com o manual do WaveWare. Contudo, foi dada especial atenção à configuração da transmissão através de MIDI dos sinais captados pelo WaveWare, uma vez que esta funcionalidade é essencial para a comunicação com o PD (consultar a secção 5.4 para mais detalhes sobre a configuração da transmissão MIDI do WaveWare e ligação ao PD).

Na Figura 5.18 mostra-se o WaveWare já configurado e a captar em tempo real os três sinais fisiológicos: em cima o GSR, no meio o HR e em baixo o muscular.

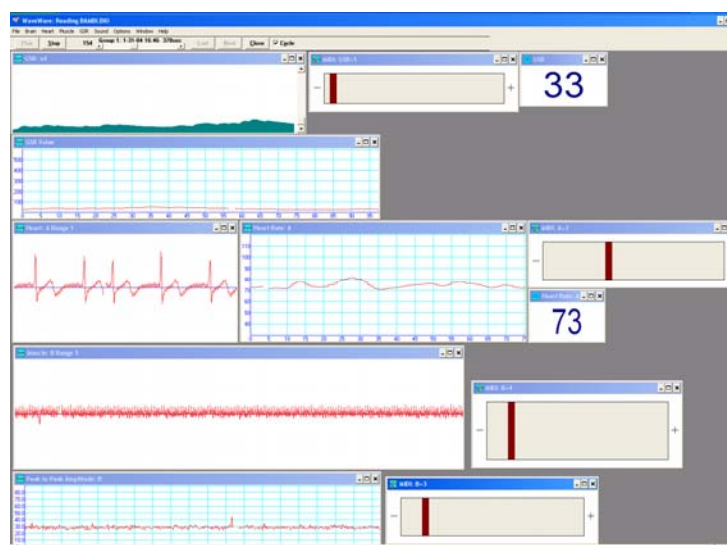


Figura 5.18 – WaveWare a captar em simultâneo o sinal HR, GSR e muscular

### 5.3.2 Interligação entre o Módulo de Apresentação e o WaveRider

Para fazer a transmissão dos dados fisiológicos entre o WaveWare e o PD foi utilizado o protocolo MIDI. O WaveWare permite que em tempo real todos os sinais captados possam ser transmitidos através de MIDI, sendo possível associar cada sinal a um canal MIDI à escolha, entre os 16 disponíveis no protocolo.

A entrada MIDI do PD apenas se torna activa quando existe uma interface externa, em *hardware*, capaz de receber ou enviar dados MIDI, ou quando existe uma interface virtual interna, em *software*, que simule uma interface em *hardware*, os chamados *patch drivers*. O MIDI Yoke NT é um dos vários softwares que oferece essa funcionalidade (MIDIOW, 2003). Fazendo uma analogia, o MIDI Yoke NT serve como um “cabo de ligação MIDI” entre a saída de um programa e a entrada de outro, neste caso o WaveWare e o PD. Para então se fazer a captação dos dados MIDI de saída do WaveWare utilizou-se o *patch driver* MIDI Yoke NT para captar a saída do WaveWare e direccioná-la para a entrada MIDI do PD.

O esquema da comunicação entre o WaveWare e o PD através da utilização do MIDI Yoke NT é o seguinte:

WaveWare -> Out 1 MidiYoke NT In 1 -> PD

O MIDI Yoke NT disponibiliza 8 portos, possuindo cada um destes portos 16 canais MIDI. No total, tem-se assim 128 canais MIDI disponíveis para comunicação. Em termos práticos, para que o envio da informação MIDI do WaveWare através do MIDI Yoke NT fosse possível, teve que se ter em atenção à configuração do menu Sound e à opção Change MIDI Device do WaveWare. Nesta caixa de diálogo escolheu-se MIDI Yoke NT 1 e o porto 1 como sendo a saída para a qual os dados captados seriam enviados, possuindo este porto 16 canais para o efeito.

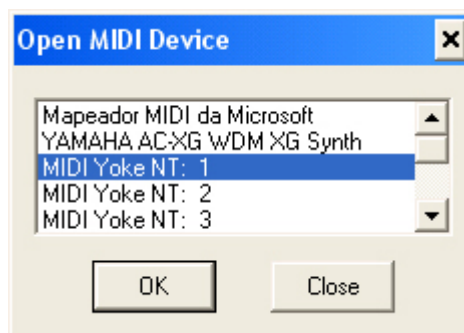


Figura 5.19 – Caixa de diálogo do WaveWare para a selecção do dispositivo MIDI de saída

Depois foi necessário definir, entre os 16 canais disponíveis nesse porto, o canal no qual cada um dos sinais captados iria ser transmitido. Para tal, acedeu-se ao Menu Sound - Open MIDI Output e abriram-se 3 janelas chamadas “MIDI”, uma para cada um dos sinais.

Em cada uma dessas janelas é possível definir o mapeamento dos sinais (GSR, HR e muscular) para uma das 16 saídas MIDI disponíveis. Através de um duplo clique, aparece a janela MIDI Options para cada um dos sinais, na qual se pode fazer esse mapeamento. No sistema implementado escolheu-se GSR no canal 1, HR no canal 2 e muscular no canal 3.

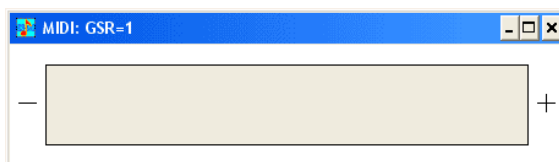


Figura 5.20 – Caixa de diálogo da saída MIDI para um sinal, neste caso o sinal GSR no canal MIDI 1

Na Figura 5.21 apresenta-se como exemplo a janela MIDI Options para o HR e a configuração utilizada para transmitir este sinal no canal 2.

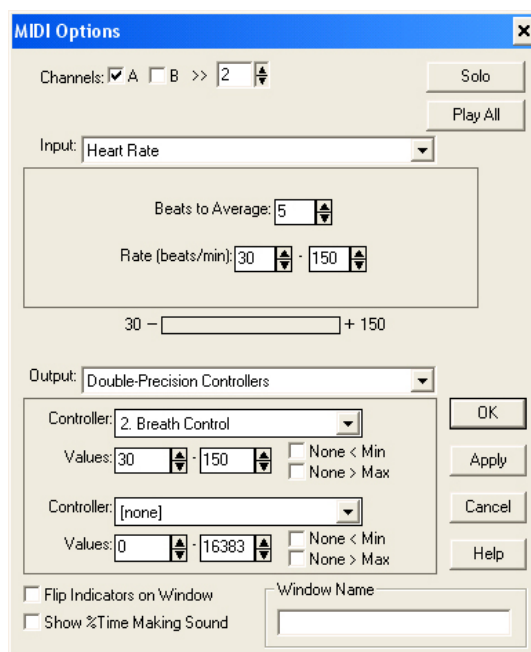


Figura 5.21 – Opções de transmissão MIDI para o sinal captado, neste caso, canal MIDI 2 para o HR

Note-se que na definição da saída MIDI, em Output, e para cada um dos 3 sinais, se escolheu o campo “Double Precision Controllers” (valores com uma precisão de 14 bits), de modo a que a representação dos valores dos dados transmitidos em MIDI fossem exactamente iguais aos que foram captados. Os outros tipos de saídas disponíveis, “Notes and Velocity” e “Controllers”, possuem apenas uma gama de valores inteiros entre 0 e 127. Uma vez que os valores dos sinais captados podem alcançar valores além do limite máximo de 127, a conversão da escala original dos sinais captados para esta adulteraria os valores dos sinais originais quando fossem transmitidos.

Depois de configurado o WaveWare para a transmissão MIDI, torna-se necessário configurar o PD para que possa receber essa mesma informação. É no ficheiro *batch* de abertura do PD (ficheiro com comandos para o carregamento de bibliotecas e de objectos adicionais ao PD, como o GEM), que é possível definir qual o porto de entrada MIDI do programa a ser utilizado, neste caso o porto 1 do MIDI Yoke NT. O comando que é necessário adicionar ao ficheiro *batch* é o seguinte:

pd -midiindex 1 (...)

A selecção do canal no qual se pretende receber a informação é feita pelo objecto [midiin <número do canal>], o qual devolve na sua saída os valores MIDI transmitidos nesse canal. O PD passa assim a poder receber informação MIDI transmitida nos 16 canais do porto 1.

```

D:\pd\bin>pd -listdev
MIDI input device #1: MIDI Yoke NT: 1
MIDI input device #2: MIDI Yoke NT: 2
MIDI input device #3: MIDI Yoke NT: 3
MIDI input device #4: MIDI Yoke NT: 4
MIDI input device #5: MIDI Yoke NT: 5
MIDI input device #6: MIDI Yoke NT: 6
MIDI input device #7: MIDI Yoke NT: 7
MIDI input device #8: MIDI Yoke NT: 8
MIDI output device #1: YAMAHA AC-XG WDM XG Synth
MIDI output device #2: MIDI Yoke NT: 1
MIDI output device #3: MIDI Yoke NT: 2
MIDI output device #4: MIDI Yoke NT: 3
MIDI output device #5: MIDI Yoke NT: 4
MIDI output device #6: MIDI Yoke NT: 5
MIDI output device #7: MIDI Yoke NT: 6
MIDI output device #8: MIDI Yoke NT: 7
MIDI output device #9: MIDI Yoke NT: 8
MIDI output device #10: Sintetizador por software GS Wa
audio input device #1: YAMAHA AC-XG WDM Audio
audio output device #1: YAMAHA AC-XG WDM Audio
not using MIDI input (use 'pd -midiindev 1' to override)
D:/pd

```

Figura 5.22 – Comando 'pd -listdev' onde é possível visualizar as entradas e saídas MIDI disponibilizadas no PD

## 5.4 Módulo de Reconhecimento de Emoções

A implementação do Módulo de Reconhecimento de Emoções passou por várias fases, fases estas que resultaram na criação de uma rede neuronal capaz de detectar o estado emotivo de um indivíduo tendo por base 3 sinais fisiológicos: HR, actividade muscular e GSR.

As fases para a implementação deste módulo foram as seguintes:

- Recolha de dados fisiológicos de um indivíduo através da realização de várias sessões para esse efeito;
- Processamento em Matlab dos dados recolhidos e treino de uma rede neuronal;
- Implementação em tempo-real do sistema de reconhecimento de emoções.

Descreve-se de seguida cada uma destas fases de uma forma mais detalhada.



### **5.4.1 Recolha de dados fisiológicos**

A recolha dos dados fisiológicos de um indivíduo foi feita com base no trabalho de James J. Gross e Robert W. Levenson, do Departamento de Psicologia da Universidade da Califórnia (Gross, 1995).

Neste trabalho de 5 anos pretendeu-se encontrar um conjunto de excertos de filmes que pudessem despertar 8 estados emotivos (alegria, raiva, bem estar, repulsa, medo, neutro, tristeza e surpresa). Para esse trabalho foram seleccionados mais de 250 excertos de filmes os quais foram apresentados a um conjunto de 494 pessoas de diversas etnias, tendo em comum o facto de falarem a língua Inglesa.

Desse conjunto inicial de mais de 250 excertos de filmes, escolheram-se os 2 melhores para cada uma das emoções que se pretendia despertar, tendo essa escolha sido feita atendendo às respostas a um questionário dado a preencher aos indivíduos depois da visualização dos excertos (Gross, 1995).

O resultado do trabalho de James J. Gross e Robert W. Levenson mostram efectivamente que os 2 filmes seleccionados para cada uma das emoções são capazes de despertar as emoções pretendidas. As emoções alvo no âmbito deste protótipo são em número mais reduzido quando comparadas com as 8 emoções do trabalho desenvolvido referido, sendo essas emoções a alegria, tristeza, susto e repouso. Para despertar as emoções pretendidas foram adquiridos os excertos dos filmes seleccionados por James Gross e Robert Levenson para despertar as mesmas emoções.

Esses excertos são os seguintes:

- Alegria – When Harry Meets Sally;
- Tristeza – Bambi;
- Susto – Capricorn One.

Para o estado de repouso não se utilizou nenhum excerto de vídeo, optando-se antes por um CD áudio de meditação.

A metodologia utilizada para a recolha dos dados fisiológicos relativos a cada estado emotivo consistiu na realização de várias sessões, uma para cada estado emotivo, divididas em duas partes distintas (Lessa, a publicar).

A captação de um sinal fisiológico resultante de um estímulo, para que o sinal possa ser efectivamente detectado e recolhido para estudo, deve ser feita em relação a um valor de base, o chamado valor basal dos sinais fisiológicos do indivíduo, obtido através de repouso ou de meditação. O valor basal permite assim que os valores obtidos num estímulo possam ser comparados com valores de referência conhecidos (esse mesmo valor basal). Quando é atingido o valor basal (através de repouso ou meditação) poderá aplicar-se um estímulo previamente catalogado como produtor de um certo estado emotivo e fazer-se a captação dos sinais durante esse estímulo.

Foram feitas 3 sessões para a recolha dos dados respeitantes aos estímulos da alegria, tristeza e susto. Na primeira parte de cada sessão, a pessoa que se disponibilizou para a recolha dos dados fisiológicos esteve em repouso durante 23 minutos, ao som de um CD de meditação. Durante esse período de repouso foi feita a captação do seu valor basal. Na segunda parte de cada sessão, depois do período de repouso, foi aplicado um estímulo visual e sonoro para a captação do sinal resultante do estímulo. Os estímulos para as emoções que se pretendiam captar foram conseguidos através dos excertos dos filmes indicados no artigo já referido (Lessa, a publicar).

Quer durante o período de repouso quer durante o período do estímulo, foram gravados os 3 sinais fisiológicos, HR, GSR e sinal muscular. No final das sessões de recolha de dados obtiveram-se no total 6 ficheiros gravados pelo WaveWare: 2 ficheiros para cada uma das 3 sessões, um relativo à fase de pré-estímulo, outro relativo à fase de estímulo.

Depois de efectuadas as recolhas, foi feita uma exportação de todos os ficheiros de dados das sessões de recolha para um ficheiro de texto (\*.txt), através da opção do WaveWare para exportação de dados: "Ascii File Output". Uma vez que o formato original dos dados é um formato proprietário do WaveWare (\*.bio), houve necessidade de fazer essa exportação para que os dados pudessem ser lidos e analisados em Matlab. Apresenta-se de seguida um exemplo do resultado dessa exportação, durante 1 segundo de captação dos sinais:

7.000,24.10,27.29,89.70  
7.125,24.10,28.25,89.70  
7.250,24.10,26.33,89.70  
7.375,24.10,28.25,89.70  
7.500,24.10,27.29,89.70  
7.625,24.10,32.07,90.40  
7.750,24.10,32.07,90.40  
7.875,24.10,27.29,90.40  
8.000,23.29,27.29,90.40

Os dados estão organizados em 4 colunas que representam as seguintes variáveis: o tempo em milésimos de segundo, o valor de GSR, o valor do sinal muscular e o valor do batimento cardíaco (HR). O WaveWare capta os sinais a uma frequência de 4Hz, daí o valor do tempo ser incrementado 125ms, ou seja, durante 1 segundo de captação de dados são recolhidas 8 amostras para cada um dos sinais.

Depois de obtidos cada um destes ficheiros de texto, a partir dos dados originais, procedeu-se à sua análise em Matlab.

#### **5.4.2 Tratamento dos dados em Matlab e redes neuronais**

Utilizou-se o Matlab e uma implementação de redes neuronais para o reconhecimento das emoções com base nos dados recolhidos pelo WaveWare.

Fui utilizada uma implementação de uma rede neuronal feed-forward backpropagation, uma vez que estas são as mais utilizadas em processos de *data mining* e reconhecimento de padrões, que é o que se pretende com os dados em causa.

Para a criação de uma rede neuronal feed-forward backpropagation de N camadas, o Matlab utiliza a função `newff`, a qual define a estrutura da rede, o tamanho e o tipo de algoritmo de treino, inicializando-a automaticamente. As últimas 2 letras da função `newff` indica o tipo de rede neuronal em causa, neste caso uma rede feedforward (MathWorks, 2004b). Os argumentos dessa função são os seguintes:

```
net = network(numInputs, numLayers, biasConnect, inputConnect, layerConnect,  
outputConnect, targetConnect)
```

```
net = newff(PR, [S1 S2...SN1], {TF1 TF2...TFN1}, BTF, BLF, PF)
```

*PR* – matriz  $R \times 2$  de valores mínimos e máximos para *R* elementos de entrada

*Si* – tamanho da *i* layer, para *N1* camadas

*TFi* – função de transferência da camada *i*

*BTF* – função de aprendizagem de Backpropagation da rede

*BLF* – função de aprendizagem do peso/desvio da rede

*PF* – função de performance

Esta função cria uma rede neuronal cuja primeira camada possui pesos vindos da entrada. Cada camada seguinte tem um peso vindo da camada anterior, tendo todas as camadas um desvio. A última camada é a saída da rede.

Foi implementada uma rede neuronal de 3 camadas, com a função de transferência de logsig para o reconhecimento de 4 estados emotivos (além dos estados emotivos de alegria, tristeza e susto, é ainda necessária a detecção do estado de repouso). A camada 1 da rede é composta por 10 neurónios, a camada 2 por 5 e a camada 3 por 4 (uma vez que era este o número de estados emotivos que se pretendiam detectar).

O número de neurónios para cada uma das camadas é definido em função dos valores de entrada através da variável *N\_IC*, relacionada com o tamanho desses dados de entrada:

- camada 1:  $S1 = N\_IC * (2 * N\_IC + 1)$ ;
- camada 2:  $S2 = 2 * N\_IC + 1$ ;
- camada 3:  $S3 = 4$ ;

O número de neurónios da camada 3 foi fixado para 4 uma vez que era este o número de estados emotivos que se pretendiam detectar.

Mostra-se de seguida parte do código em Matlab para a definição e criação da rede:

```
function [err,net,transMat]=treina_uma_rede_neuronal(mydata,class)
    [c1,c2,c3,c4]=gera_classes_binarias_mesmo(class);
    L=length(class);
    c=zeros(4,L);
    c(1,:)=c1;
```

```

c(2,:)=c2;
c(3,:)=c3;
c(4,:)=c4;

IC_ori=mydata';
[N_IC_ori,lixo]=size(IC_ori);
[IC,transMat] = prepca(IC_ori,0.01);

[N_IC,lixo]=size(IC);

% - definição do número de neurónios para cada uma das 3 camadas
% -----
S1= N_IC*(2*N_IC+1); S2=2*N_IC+1; S3= 4;
% -----

% criação da rede neuronal
% -----
net=newff(minmax(IC),[S1 S2 S3],{'logsig' 'logsig' 'logsig'},'trainlm');
% -----
net.initFcn='initlay';

for i=1:length(net.layers)
    net.layers{i}.initFcn='initwb';
end;
[a,b]=size(net.inputWeights);
for i=1:a
    for j=1:b
        net.inputWeights{i,j}.initFcn='rands';
        net.layerWeights{i,j}.initFcn='rands';
    end;
end;
for i=1:length(net.biases)
    net.biases{i}.initFcn='rands';
end;

net.performFcn='mse';
net.trainParam.goal=0.01;
net.trainParam.show=1;
net.trainParam.epochs=500;
net.trainParam.mc=0.95;
net.trainParam.time=inf;
net.trainParam.mu_inc=5;

net = init(net);
[net,tr]=train(net,IC,c);

y=transMat*IC_ori;
out=sim(net,y);
for i=1:length(out)
    [vmax,pmax]=max(out(:,i));

```

```

        resposta(i)=pmax;
    end;
    err=sum(abs(class-resposta)~=0)/L;
return;

```

Tal como já se descreveu no Capítulo 3 - Ferramentas, o processo para a implementação e utilização de redes neuronais passa primeiro pelo treino da rede antes de esta poder ser utilizada. Esse treino foi assim efectuado tendo por base os ficheiros obtidos a partir das sessões de recolha de dados previamente efectuadas. Foram utilizados 3 ficheiros para as emoções a detectar: o ficheiro relativo ao estímulo do susto, estímulo da alegria e estímulo da tristeza. Como se pretendia detectar também o estado de repouso, utilizou-se o ficheiro obtido na primeira parte da sessão referente ao estímulo do susto, perfazendo um total de 4 ficheiros para o treino da rede. A escolha deste ficheiro relativo ao repouso deveu-se ao facto de ser este o que possuía dados basais mais uniformes e com o menor número de artefactos detectados ao longo do processo de recolha.

O treino da rede é feito pela função `train`, cuja sintaxe é a seguinte:

```
[net,tr,Y,E,Pf,Af] = train(net,P,T,Pi,Ai,VV,TV)
```

```
[net,tr]=train(net,IC,c);
```

`net` – rede neuronal

`IC` – entrada da rede

`c` – saída alvo da rede

Para o treino da rede utilizaram-se os seguintes valores para os parâmetros:

`net.performFcn='mse'`; - função utilizada para a avaliação da performance da rede, neste caso *mse* - *mean squared error*

`net.trainParam.goal=0.01`; - objectivo da performance a alcançar pela rede em termos de tolerância de erro, neste caso, inferior a 1%

`net.trainParam.show=1`; - número de iterações entre as quais se mostram os resultados do treino da rede

`net.trainParam.epochs=500`; - número máximo de iterações de treino da rede até que os objectivos sejam alcançados

`net.trainParam.mc=0.95`; - coeficiente de *momentum*, possibilita o treino da rede aquando da existência de mínimos locais nos dados;

`net.trainParam.time=inf`; - tempo máximo de treino da rede, neste caso infinito

`net.trainParam.mu_inc=5`; - parâmetro utilizado para definir a velocidade de convergência do algoritmo de treino.

Estes parâmetros condicionam todo o treino da rede e os resultados alcançados.

Depois do treino da rede com os dados das sessões de recolha realizadas, a rede em funcionamento produz um vector de inteiros entre 1 e 4, representando cada um dos valores um estado emotivo: 1 – tristeza; 2 – susto, 3 – alegria; 4 – repouso.

Como se disse, os valores de saída, inteiros entre 1 e 4, são os valores que a rede produzirá quando estiver a ser utilizada com dados obtidos em tempo real. Ou seja, os dados captados em tempo real pelo WaveRider, transmitidos ao PD e deste à rede neuronal, serão analisados pela rede, e consoante o estado emotivo detectado, o resultado dessa detecção pelo Matlab terá como resultado um *array* de valores inteiros entre 1 e 4. Posteriormente, o *array* é transmitido ao Java onde se procederá à sua análise final de modo a se poder obter desse *array* um único inteiro representativo do estado emotivo a transmitir ao PD (Ver o ponto 5.5.3 para mais detalhes do funcionamento da rede em tempo real).

Apresenta-se na Figura 5.23 um exemplo da saída da rede quando a funcionar com os dados obtidos pelo WaveWare: a rede detecta o estado 3 – alegria com sucesso, existindo contudo algumas detecções falsas em que é detectado o estado 2 – susto.

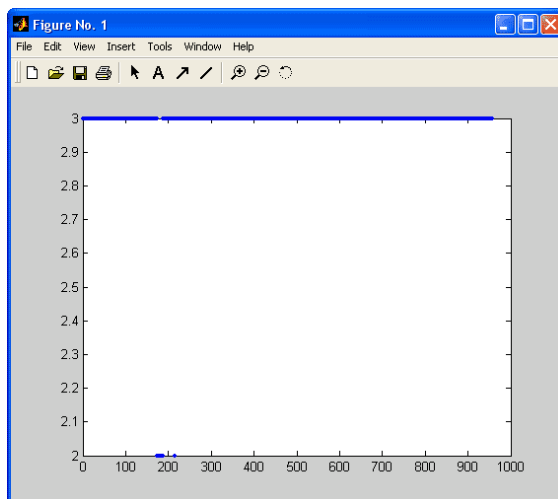


Figura 5.23 – Janela do Matlab da saída da rede neuronal, detectando o estado 3, alegria

Para utilização da rede em tempo real, é necessário proceder a uma codificação inicial dos dados de entrada da rede. Essa codificação foi

efectuada através de codificação diferencial e normalização dos dados (primeira e segunda derivada sendo depois aplicada uma normalização aos valores médios e máximos).

A função utilizada para aplicar a rede neuronal aos dados captados é a que se apresenta de seguida:

```
function resposta=aplica_rede_neuronal(net_file_name,in_data_file_
name,reference_data_file_name,modo,t)

    in_data=le_estes_dados(in_data_file_name);
    reference_data=le_estes_dados(reference_data_file_name);

    % Normaliza os dados in_data de acordo com os dados de referencia
    % contidos em reference_data:
    in_data_norm=codifica_usando_referencia(in_data,reference_data);
    in_data_norm=in_data_norm(:,1:3*modo);

    mydata=rolling_data(in_data_norm,t);

    % Vai ler a rede neuronal e transformacao de pre-processamento ao
    % ficheiro file_name:
    [TransMat,net]=le_esta_rede_neuronal(net_file_name);

    y=TransMat*mydata';
    out=sim(net,y);
    for i=1:length(out)
        [vmax,pmax]=max(out(:,i));
        resposta(i)=pmax;
    end;
return;

function out=le_estes_dados(file_name)
    in_data=load(file_name);
    f_name=fieldnames(in_data);
    out=getfield(in_data,f_name{1,1});
    out=out(:,2:4);
return;

function [TransMat,net]=le_esta_rede_neuronal(net_file_name)
    in_data=load(net_file_name);
    f_name=fieldnames(in_data);
    TransMat=getfield(in_data,f_name{1,1});
    net=getfield(in_data,f_name{2,1});

return;
```



A função utilizada pelo Matlab, para aplicar a rede neuronal já treinada em dados obtidos em tempo real é a seguinte:

```
out=sim(net,y);
```

net – a rede neuronal

y – entrada da rede

out – a saída da rede

### **5.4.3 Implementação em tempo-real do sistema de reconhecimento de emoções**

O processo de reconhecimento das emoções é controlado pelo PD, que, de 20 em 20 segundos, inicia esse processo através de um temporizador. Este intervalo de 20 segundos é facilmente ajustável e facilmente alterado, mas considerou-se este valor uma vez que:

- valores mais baixos poderiam despoletar estados erróneos devido ao possível ruído do sinal (por exemplo, caso a pessoa se tenha mexido durante um período de repouso o súbito aumento do sinal muscular poderia levar a interpretação do estado como de susto). Assim, quanto maior o tempo de recolha de sinais, menos influência o ruído tem;
- valores mais altos detectariam com mais fiabilidade os estados emotivos, mas emoções de resposta rápida (tal como susto), caso fossem detectadas, demorariam um longo período de tempo a serem apresentadas, o que para um protótipo desenvolvido para funcionar em tempo-real não seria razoável.

Considerou-se assim o valor de 20 segundos como um valor intermédio aceitável para permitir uma boa captura de sinais, de modo a evitar o possível ruído dos sinais, e suficientemente rápido para se poder obter uma resposta do sistema em tempo-real.

Para se interligar o PD como a rede neuronal em Matlab utilizou-se a linguagem Java. Os programas interligam-se sendo o PD cliente do Java e o Java por sua vez cliente do Matlab, utilizando-se para o efeito a biblioteca JMatLink para efectuar a interligação entre Java e Matlab (Muller, 2004).

No outro sentido, o Matlab é servidor do Java e o Java por sua vez do PD. O PD possui um *external*, o *maxlib*, que inclui os objectos [netserver] e [netclient] construídos por Olaf Matthes para ligação a um *host* local ou remoto e permutar informação, na base de cliente-servidor, podendo o PD ser cliente ou servidor. Na Figura 5.24 está representada a troca de informação entre as aplicações, quer como cliente quer como servidor, representando a estrutura da implementação uma arquitectura *three-tier*.

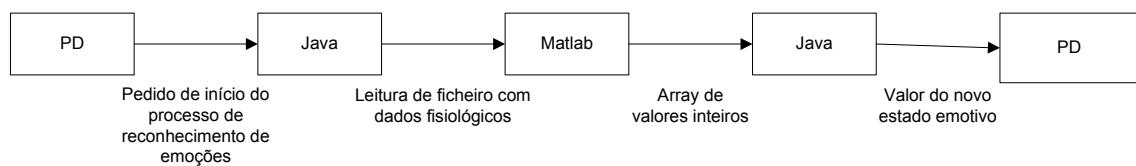


Figura 5.24 – Interligação e passagem de informação entre *packages* de desenvolvimento diferentes. Arquitectura cliente-servidor ou *three-tier*

O Java desempenha o papel de *middleware* em sistemas de computação, que significa que funciona como um agente de software actuando como um intermediário, ou como um membro de um grupo de intermediários entre diferentes componentes num processo transaccional.

Em PD foi implementada uma abstracção “matbrain” que inclui os *patches* responsáveis pelo reconhecimento das emoções com base nos sinais MIDI recebidos e pelo registo do histórico do sistema numa base de dados, sendo esses *patches*:

- *patch* brain – *patch* que recebe os dados MIDI e activa o programa em Java responsável pela detecção do estado emotivo com base nesses dados;
- *patch* sub\_brain\_alterado – *patch* responsável pela activação do Pacote correspondente ao valor enviado pelo programa em Java após o reconhecimento da emoção;
- *patch* write\_pd – *patch* responsável pela escrita na base de dados do histórico do sistema, ou seja, pelo registo temporal dos estados emotivos detectados.

Na Figura 5.25 apresenta-se a abstracção mat\_brain e os *patches* que a constituem.

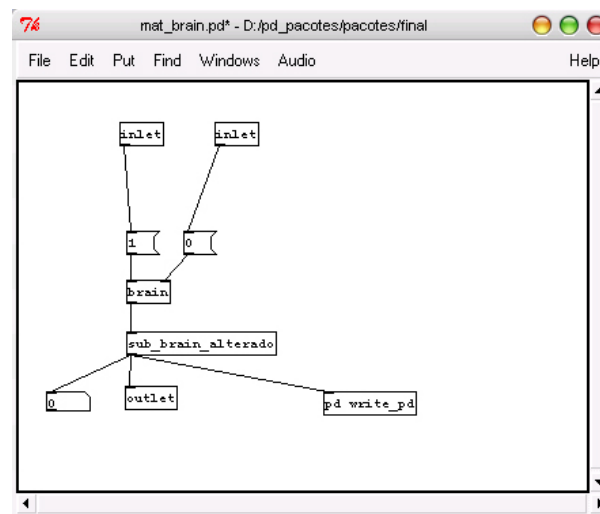


Figura 5.25 – Patch mat\_brain.pd

O processo de detecção e registo dos estados emotivos é controlado pelo PD, uma vez que é este que despoleta todas as acções que desencadeiam o processo de detecção e registo da informação.

Todo o processo tem início no WaveWare, o software de recolha dos sinais fisiológicos em tempo-real. O WaveWare está continuamente a enviar os dados do HR, GSR e sinal muscular para o PD através do protocolo MIDI. O PD, através do *patch* brain, recebe esses dados e de 20 em 20 segundos grava-os para um ficheiro de texto num formato igual ao obtido aquando da exportação dos ficheiros \*.bio do WaveWare para ficheiros de texto \*.txt.

A recepção dos dados MIDI enviados pelo WaveWare é feita pelo objecto [ctlin <número do canal MIDI>], estando a escrita desses mesmos dados no ficheiro dadospd.txt a cargo do objecto [textfile]. Apresenta-se na Figura 5.26 parte do *patch* responsável pela recepção e escrita dos dados MIDI recebidos.

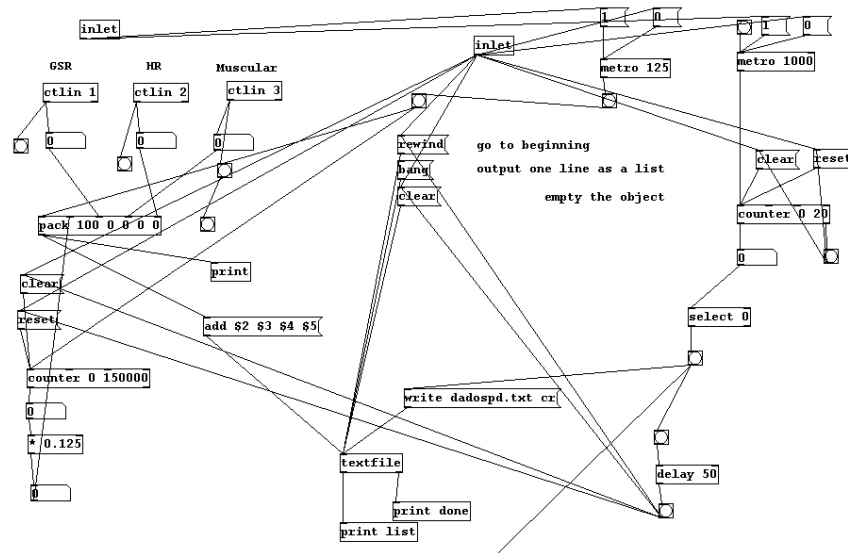


Figura 5.26 – Parte do *patch* brain.pd responsável pela recepção e escrita dos dados MIDI enviados pelo WaveWare

Depois de gravado o ficheiro dadospd.txt com a informação dos sinais dos últimos 20 segundos, o PD envia uma mensagem para o programa Client.java de modo a que a análise para a detecção do estado emotivo possa ter início. O envio desta mensagem é feito através do objecto [netserver 8002], o qual cria uma ligação através de *sockets* com o programa em Java (Java, 2004b). Este objecto possibilita o envio e a recepção de informação entre o PD e o programa Java.

A criação em Java do cliente capaz de receber e enviar informação através do porto 8002 utilizando *sockets* faz-se através do seguinte código:

```
Socket kkSocket = null;
PrintWriter out = null;
BufferedReader in = null;
double[][] pre_classe;
double[][] array2d; // = new double[4][4];

JMatLink engine = new JMatLink();

try {
    kkSocket = new Socket("localhost", 8002);
    out = new PrintWriter(kkSocket.getOutputStream(), true);
    in = new BufferedReader(new InputStreamReader(kkSocket.getInputStream()));
} catch (UnknownHostException e) {
    System.err.println("Don't know about host: localhost.");
    System.exit(1);
} catch (IOException e) {
    System.err.println("Couldn't get I/O for the connection to: localhost.");
    System.exit(1);
}

BufferedReader stdIn = new BufferedReader(new InputStreamReader(System.in));
String fromServer;
String fromUser;
int count = 0;
```

O programa Java está continuamente à espera da mensagem de controlo enviada pelo PD através de um *socket* para que possa dar início ao reconhecimento das emoções com base nos dados escritos no ficheiro de texto. A mensagem de controlo enviada pelo PD é simplesmente um inteiro, “777”, que quando enviado, despoleta em Java o processo de reconhecimento das emoções.

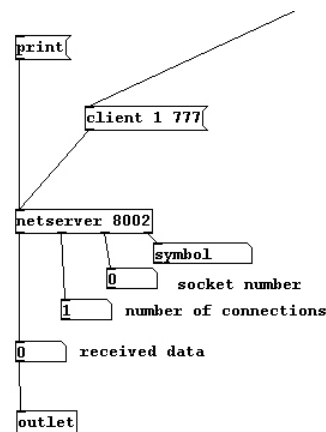


Figura 5.27 – Parte do *patch* brain.pd responsável pela análise feita em Java e Matlab. O resultado obtido dessa análise é recebido no primeiro *outlet*, *received data*

Através da biblioteca JMatLink é possível interligar o Java com o Matlab através de métodos nativos, tornando-se possível utilizar o poder computacional do Matlab dentro de um programa Java (Muller, 2004). Esta biblioteca possibilita assim que se possam escrever em Java os comandos a enviar ao Matlab, da mesma forma como se escreveriam dentro do próprio Matlab.

Os comandos Matlab para que se possa fazer o reconhecimento das emoções são os seguintes:

```
engine.engEvalString("a = dlmread('d:\pd_Pacotes\\Pacotes\\final\\dadospd.txt',' ');");
engine.engEvalString("save data.mat a;");
engine.engEvalString("load data.mat;");

engine.engEvalString("pre classe =
aplica_rede_neuronal('rede_neuronal.mat','data.mat','data_capricorne_cd',1,6);");

pre_classe = engine.engGetArray("pre_classe");
```

Os dados do ficheiro de texto dadospd.txt criado pelo PD são lidos pelo Matlab e gravados no seu formato nativo. Depois de guardado o ficheiro, aplica-se a rede neuronal sobre esse ficheiro. Como resultado é retornado um *array* *pre\_classe* com os valores do estado

emotivo detectado. A análise final deste *array* é feita em Java para a detecção do estado emotivo.

Como no PD é necessário receber um único inteiro de 20 em 20 segundos, representativo do estado emotivo (de modo a que se possa mostrar o módulo da emoção adequado), é feita em Java, a cada 20 segundos, uma análise do *array* obtido no Matlab e na rede neuronal e enviado o resultado para o PD.

Esta análise em Java baseia-se na contagem de cada um dos valores inteiros de 1 a 4 presentes no *array* e na obtenção da percentagem de cada um desses valores. Considerou-se o valor de 90% de determinado inteiro como sendo uma boa margem para que o estado emotivo possa ser considerado fiável, isto é, caso a percentagem de um inteiro seja superior a 90% em relação aos outros inteiros presentes no *array*, considera-se que o estado emotivo é o representado por esse inteiro. Caso não se verifique essa condição é retornado um inteiro ("77") que tem a função de *flag* para o PD. Esta *flag* pretende transmitir ao PD que um novo estado emotivo não foi detectado com a precisão suficiente, mantendo-se no PD o estado anterior.

A parte do programa em Java responsável por esta análise apresenta-se de seguida:

```
int getClass (double[][] array) {

    double p_1, p_2, p_3, p_4;
    double p_valor1; p_valor1 = 0;
    double p_valor2; p_valor2 = 0;
    double p_valor3; p_valor3 = 0;
    double p_valor4; p_valor4 = 0;

    double tamanho;

    tamanho = array[0].length;

    for(int i=0; i < array[0].length; i++) {
        if (array[0][i]==1.0) {
            p_valor1 = p_valor1 +1;
        } else if (array[0][i]==2.0) {
            p_valor2 = p_valor2 +1;
        } else if (array[0][i]==3.0) {
            p_valor3 = p_valor3 +1;
        } else if (array[0][i]==4.0) {
            p_valor4 = p_valor4 +1;
        }
    }

    p_1 = (p_valor1 / tamanho) * 100;
    p_2 = (p_valor2 / tamanho) * 100;
    p_3 = (p_valor3 / tamanho) * 100;
    p_4 = (p_valor4 / tamanho) * 100;

    int saida_temp;

    if (p_1>p_2 && p_1>p_3 && p_1>p_4 && p_1 > 90.0) {
        saida_temp = 1;
    } else if (p_2>p_1 && p_2>p_3 && p_2>p_4 && p_2 > 90.0) {
        saida_temp = 2;
    } else if (p_3>p_1 && p_3>p_2 && p_3>p_4 && p_3 > 90.0) {
        saida_temp = 3;
    } else if (p_4>p_1 && p_4>p_2 && p_4>p_3 && p_4 > 90.0) {
```

```

    }
    } else {
        saida_temp = 77;
    }
    return saida_temp;
}
}

```

A recepção no PD do valor enviado pelo programa em Java é feita, como já se referiu, através do objecto [netserver 8002] e do seu primeiro *outlet*. Quando o PD recebe um inteiro entre 1 e 4, faz a apresentação do estado emotivo correspondente a esse valor, sendo a correspondência a seguinte:

- 1 – tristeza;
- 2 – susto;
- 3 – alegria;
- 4 – repouso.

O *patch* responsável pela recepção da informação enviada pelo Java e pela apresentação do Pacote correcto no PD é o *patch* *sub\_brain\_alterado*, que se apresenta na Figura 5.28.

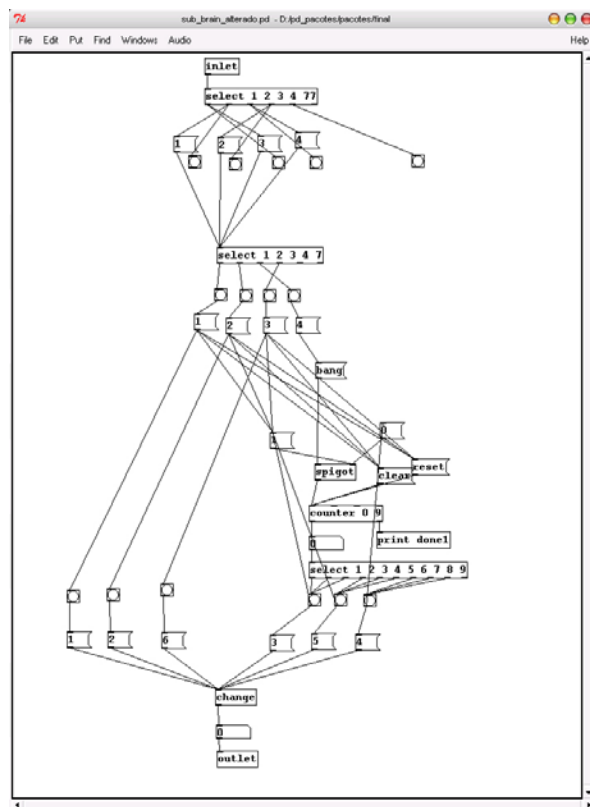


Figura 5.28 – Patch *sub\_brain\_alterado*.pd

Neste *patch*, para a selecção do Pacote a apresentar, foi dada especial atenção ao caso do Pacote da Tristeza. Tal como se referiu anteriormente, na secção 5.2.1 - Pacotes, caso a emoção detectada se mantenha por mais de duas vezes consecutivas na tristeza, é apresentado o Pacote da Tristeza do nível seguinte (o Pacote de Nível 2 caso se esteja no Nível 1, e o Nível 3 caso se esteja no Nível 2), mantendo-se no último nível (o Nível 3) até que seja detectada outra emoção. A cada um destes 3 níveis da Tristeza associou-se um inteiro para que o Pacote associado a cada um desses níveis pudesse ser referenciado e apresentado quando necessário.

Foi também dada especial atenção ao caso do reconhecimento consecutivo das emoções corresponder ao mesmo estado que o anterior. Caso essa situação se verifique, o Pacote actualmente em apresentação deverá ser mantido e não deverá ser novamente carregado até que se detecte um novo estado.

Para tal, foi utilizado o objecto [change]. Este objecto tem como função detectar e assinalar qualquer alteração do seu valor de entrada. Quando tal acontece, o seu valor de saída é igual ao novo valor de entrada, mas caso não haja alteração do valor de entrada a saída mantém-se no valor anterior. Por exemplo, quando dão entrada no objecto [change] dois valores consecutivos iguais, apenas o primeiro é retornado na saída, só voltando o objecto a produzir uma nova saída quando o valor de entrada no objecto for diferente do valor anterior nessa mesma entrada.

## **5.5 Módulo do Histórico do Sistema**

O sistema possui um histórico que regista as transições entre os estados emotivos detectados pelo sistema, incluindo a data e hora em que se deu essa transição.

O sistema de histórico foi implementado utilizando o PD, Java e a base de dados MySQL (MySQL, 2004), para o registo de informação, e PHP (PHP, 2004) e um servidor de HTTP, Apache (Apache, 2004), para a visualização dessa mesma informação.

O fluxo de informação para o registo do histórico das emoções na base de dados tem início no PD. Quando se dá uma mudança de estado, o inteiro que indica o valor do novo estado e do Pacote a ser apresentado é enviado pelo servidor de *sockets* para o cliente Java. Por sua vez, o cliente Java, quando recebe o valor grava-o na base



de dados MySQL juntamente com a data e a hora em que ocorreu essa alteração.

Para a visualização da informação foi construído um pequeno *síte* em PHP/MySQL, para que a consulta dessa informação fosse a mais prática possível.

### 5.5.1 Registo de informação

Todo o processo de registo de informação através do histórico do sistema é despoletado pelo objecto [change] no *patch* sub\_brain\_alterado quando se detecta um novo estado emotivo. Quando tal se verifica é enviado o valor desse novo estado para o *patch* write\_pd.

O *patch* write\_pd implementa um servidor de *sockets*, neste caso no porto 8004. Nesse porto 8004 encontra-se ligado um cliente implementado em Java, responsável pela recepção da informação enviada pelo PD. Após a recepção da informação enviada pelo *patch* sub\_brain\_alterado, o *patch* write\_pd envia-a imediatamente através do objecto [netserver 8004] para o cliente Java no porto 8004.

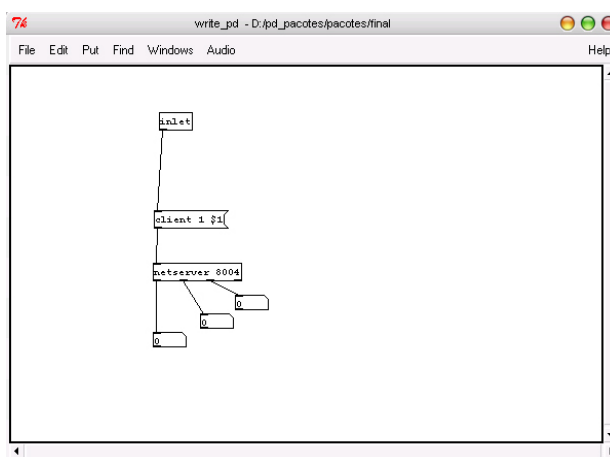


Figura 5.29 – *Patch* write\_pd.pd responsável pelo envio do valor do Pacote seleccionado no porto 8004

Este cliente Java possui uma implementação de um cliente de *sockets*. Para a criação do *socket* utilizou-se o código mostrado atrás só que agora para o porto 8004.

Após a implementação da comunicação entre o cliente e o servidor, o cliente fica à espera de receber informação através da leitura

constante da entrada no porto definido. Esta leitura é feita através de um *loop* infinito no qual o método `readLine()` faz constantemente a leitura da entrada do *socket*:

```
try {  
    for(;;) {  
        String estado = in.readLine();  
  
        (...)
```

Quando a variável “estado” recebe um novo valor dá-se início ao processo de registo na base de dados MySQL. O processo de ligação a uma base de dados através de Java envolve 2 passos: a instalação do *driver* JDBC (*Java Database Connectivity*) para a base de dados MySQL e o estabelecimento da ligação à base de dados (Java, 2004c).

O JDBC é um API (Application Programming Interface) fornecido pelo Java para que se possa comunicar com uma base de dados. Uma vez que a base de dados escolhida foi o MySQL, foi necessária a instalação de um *driver* JDBC de acesso a MySQL. O *driver* utilizado foi o [MySQL Connector/J](#), que é oficialmente o *driver* JDBC para MySQL e disponibilizado pela própria empresa que mantém o MySQL.

O carregamento do *driver* JDBC faz-se em apenas uma linha de código:

```
Class.forName("com.mysql.jdbc.Driver");
```

Uma vez o *driver* carregado, este passa a estar disponível para o estabelecimento da ligação à base de dados, o segundo passo. Este segundo passo envolve a escrita do caminho (ou url – *uniform resource locator*) que permite localizar a base de dados. Uma vez que a base de dados é MySQL, este url começa com a *string* “jdbc:mysql”, sendo o resto do caminho o “local” onde a base de dados se encontra, neste caso “//localhost:3306/from\_within”. Depois de especificado o url, o método `DriverManager.getConnection` estabelece a ligação à base de dados. Este método possui 3 argumentos: o url já especificado, o *login* e a *password* de acesso à

base de dados. Assim, o código completo utilizado para o estabelecimento da ligação é o seguinte:

```
String url = "jdbc:mysql://localhost:3306/from_within";  
Connection con = DriverManager.getConnection (url, "root", "");
```

O método `DriverManager.getConnection` retorna uma ligação aberta para a comunicação com a base de dados. Estando essa ligação aberta, neste caso a ligação “con”, podem ser utilizados comandos JDBC que incluem comandos SQL (*Structured Query Language*) para operações sobre a base de dados em causa.

O `Statement` é o objecto encarregue de enviar os comandos SQL para a base de dados. Para isso, cria-se uma instância desse objecto e executa-se, escolhendo-se o método de execução apropriado consoante o comando SQL que se quer enviar para a base de dados.

No caso do sistema de registo do histórico, necessita-se de fazer uma inserção de informação na base de dados respeitante ao novo estado emotivo detectado. Essa informação inclui a data e a hora em que essa alteração de estado se deu, assim como o próprio estado. Para tal, criou-se na base de dados uma tabela com os seguintes campos:

```
CREATE TABLE `logs` (  
  `id` int(11) NOT NULL auto_increment,  
  `ano` int(11) NOT NULL default '0',  
  `mes` int(11) NOT NULL default '0',  
  `dia` int(11) NOT NULL default '0',  
  `hora` text NOT NULL,  
  `estado` int(11) NOT NULL default '0',  
  PRIMARY KEY (`id`)  
) TYPE=MyISAM AUTO_INCREMENT=10 ;
```

O comando SQL para o registo da informação é o seguinte:

```
String query = "INSERT into logs(ano,mes,dia,hora,estado) values('"
               + year + "',''"
               + month + "',''"
               + dayOfMonth + "',''"
               + hour + ":" + minute + "',''"
               + estado + "')";
```

Este comando SQL é passado para a base de dados através do método `executeUpdate`, que tem como argumento o comando SQL:

```
try {

    Connection con = DriverManager.getConnection (url, "root", "");
    checkForWarning (con.getWarnings ());
    DatabaseMetaData dma = con.getMetaData ();
    Statement stmt = con.createStatement ();
    stmt.executeUpdate (query);
    stmt.close();
    con.close();

}
```

### 5.5.2 Leitura de informação

A leitura e apresentação da informação é feita através do servidor de HTTP Apache, no qual foi instalado o PHP 4 para a implementação de um pequeno *site* de consulta da informação do histórico do sistema. O PHP é utilizado para fazer consultas à base de dados e para a construção de páginas dinâmicas, consoante a informação contida nessa mesma base de dados.

A informação no *site* está estruturada em anos. Quando se escolhe um determinado ano é apresentada uma lista dos meses em que houve algum registo. Clicando em determinado mês, aparecem os dias desse mês em que foi registada alguma actividade emotiva. Escolhendo um dia desse mesmo mês, aparece uma listagem de todos os estados emotivos que foram registados nesse dia, com informação da hora.

A Figura 5.30 apresenta um exemplo de uma sequência de consulta da informação do histórico do sistema: consulta do ano, mês dia e emoções registadas.

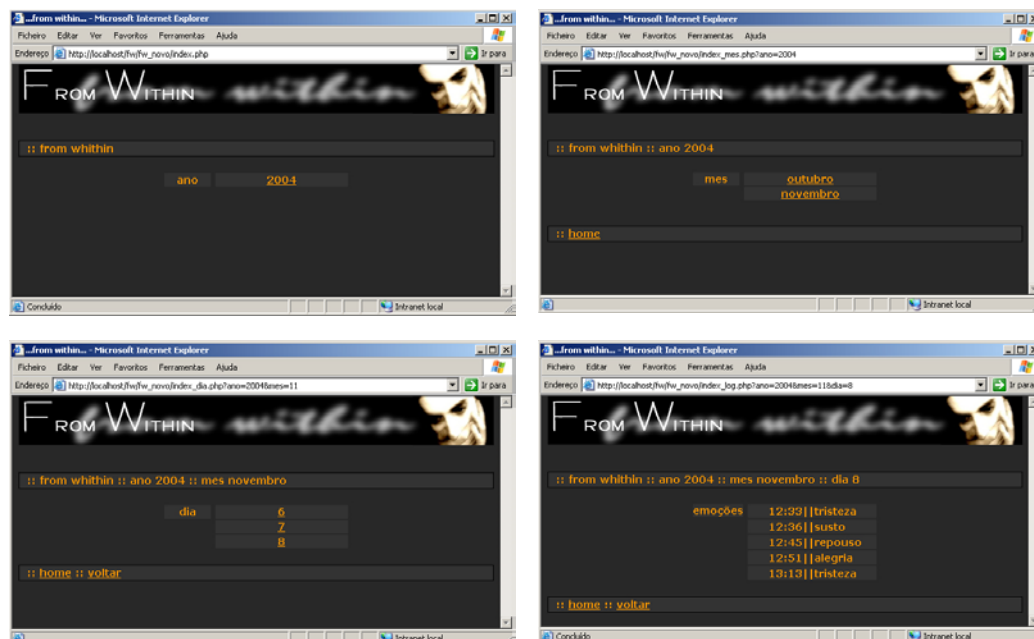


Figura 5.30 – Histórico do sistema.

## 5.6 Wearable afetivo

Para testar o protótipo a funcionar em tempo real foi construído um *wearable* em espuma. Esse *wearable* possui 3 peças: um capacete, um casaco e umas calças (Lessa, a publicar).

No capacete foi acoplada a web cam que serve de sinal de entrada de vídeo em tempo real do sistema. Possui também uma abertura frontal na qual a cara do utilizador fica completamente visível.

O casaco possui uma abertura na zona do peito para um monitor TFT, servindo este de apresentação visual do sistema. O monitor TFT é utilizado pela pessoa através de um sistema de alças que passam pelos ombros, para que este fique fixo ao corpo. O monitor deverá ser “vestido” antes do casaco, uma vez que o casaco fica por cima do monitor. Na zona dos ombros existem dois pequenos altifalantes acoplados ao casaco para a reprodução sonora do sistema.

As calças não possuem nenhuma característica em especial (Lessa, a publicar).

O *wearable* construído apresenta-se na Figura 5.3.1 (Lessa, a publicar).

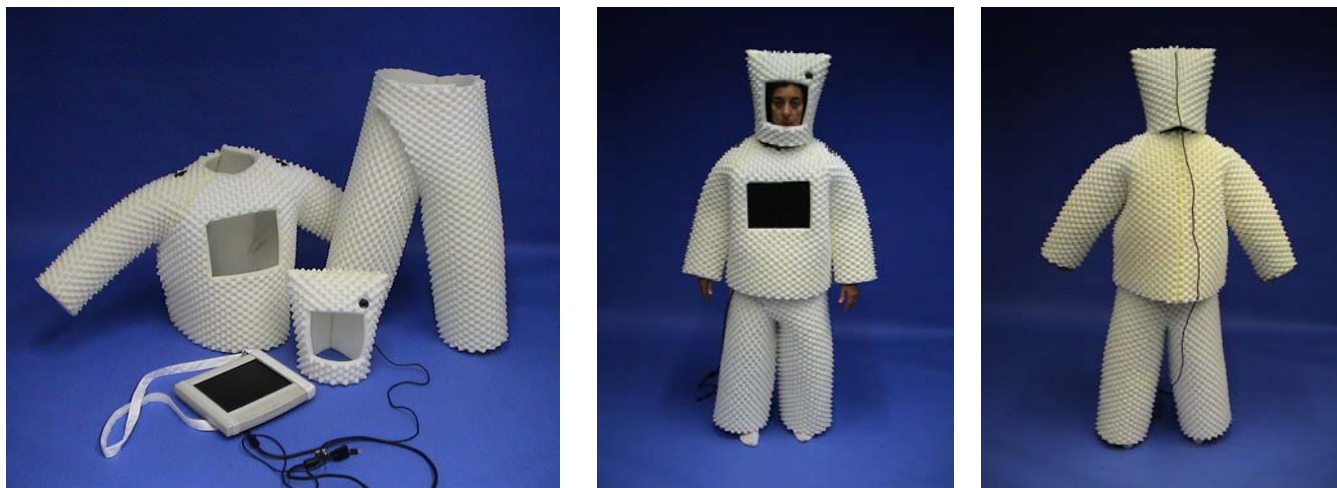


Figura 5.31 – *Wearable* afectivo (Lessa, a publicar)

Antes de se poder experimentar em tempo-real o protótipo e do utilizador vestir o *wearable*, é necessário que os sensores responsáveis pela captação dos 3 sinais fisiológicos sejam colocados no corpo do utilizador. Depois de colocados pode-se então vestir o *wearable*.

Após a colocação dos eléctrodos e estando o *wearable* vestido, pode-se ligar o monitor TFT do *wearable* à saída de um segundo monitor do computador encarregue de fazer a apresentação dos Pacotes, ligando-se igualmente a *web cam* do *wearable* a este computador. Os eléctrodos ligam-se por sua vez ao WaveRider Jr, que está ligado ao mesmo computador responsável pela apresentação visual dos Pacotes e pela execução da rede neuronal.

Uma vez que para a reprodução sonora se utiliza um outro computador, as colunas do *wearable* deverão estar ligadas a esse.

## 5.7 Inquéritos de opinião

Para se verificar até que ponto a atribuição do sentido afectivo, presente aquando da concepção dos Pacotes, encontra eco na generalidade das pessoas, procedeu-se a realização de inquéritos de opinião (Lessa, a publicar). Embora durante o projecto de implementação do Módulo de Apresentação e dos Pacotes que o constituem se teve por meta elaborar um conjunto de valores imagéticos e sonoros, os quais, no seu conjunto orientassem o espectador das sequências finais para um determinado ambiente

emocional, o âmbito de tais atribuições de imagens e sons a emoções é eminentemente subjectivo, sendo afectado por diversos factores que divergem de pessoa para pessoa.

Para a realização dos inquéritos, foram reunidos 12 voluntários, numa amostragem de 8 mulheres e 4 homens, com idades compreendidas entre 24 e os 56 anos. O inquérito consistiu no visionamento, individual, de um documento vídeo com cerca de 9 minutos, preparado para o efeito e dois questionários. O primeiro questionário foi fornecido após o visionamento do vídeo e o segundo após uma complementar informação dos excertos dos poemas inseridos nos Pacotes, excertos esses que, deliberadamente, não tinham uma leitura aquando da visualização do vídeo e da apresentação do primeiro questionário (Lessa, a publicar).

O primeiro questionário, para além das questões sobre o aspecto geral do *wearable* e suas possíveis aplicações, visava essencialmente reter o entendimento emocional dos voluntários perante o impacto visual e sonoro dos Pacotes. O segundo questionário visava verificar se essa atribuição é corroborada pelos excertos dos poemas inseridos em cada um dos Pacotes ou não (uma vez que os textos implicavam uma atenção mais dirigida após o impacto inicial).

O documento vídeo faz uma apresentação do protótipo, mostrando os elementos que o constituem e mostrando uma simulação de um cenário de utilização; e por fim, mas o que se considera ser a componente que mais interessaria avaliar, apresenta as sequências multimédia que se pretende associar a um estado afectivo mais ou menos abrangente.

Ao nível da análise das sequências multimédia e consequente atribuição de significado emocional, por parte dos voluntários que realizaram o teste, apresentam-se de seguida os resultados (Lessa, a publicar).

Para o Pacote da Alegria, associado a um estado afectivo positivo, seis voluntários (50%) associaram-no ao estado positivo; dois (16,67%) indicaram um estado híbrido de relaxamento e alegria; três (25%) consideraram-no representativo de um estado de relaxamento e por fim um voluntário (8,33%) associa-o a um estado híbrido de relaxamento e estado neutro (Lessa, a publicar).

Para o Pacote da Tristeza e as 3 Cenas que o constituem, associado a um estado afectivo negativo relacionado com nostalgia e tristeza, no qual o grau de tristeza se intensificaria da primeira à última Cena, os resultados são: sete voluntários (58,34%) identificaram-no como negativo, dos quais se distinguem quatro que conectam o conteúdo do Pacote com um sentido de introspecção e os restantes três com a ideia de “conhecer a verdade”; três voluntários (25%) associaram as duas últimas Cenas da do Pacote ao estado emotivo negativo (tristeza), dos quais se distinguem dois que identificam a 1ª Cena com relaxamento e o terceiro voluntário que a identifica com alegria; um voluntário (8,33%) relaciona a as Cenas com um estado emocional misto entre negatividade, relaxamento e alegria; um voluntário (8,33%) associa-a com um estado de ansiedade (Lessa, a publicar).

No Pacote do Susto verifica-se que: nove dos voluntários (75%) identificam-no como uma representação de um estado de ansiedade, complementando com referências a nervosismo (1), insegurança (1), receio e medo (1), confusão (1), medo e stress (1) e irritação (1); um voluntário (8,33%) revê nesta sequência um estado de estranheza e calma; os restantes dois voluntários (16,67%) associam-na a um estado negativo, dos quais um dos dois complementa esta informação com a ideia de loucura e obsessão (Lessa, a publicar).

Finalmente para o Pacote de Repouso os resultados são: dez voluntários (83,33%) identificam-na como um estado de relaxamento e complementam a sua associação aos termos apatia (2), adormecimento (1), rotina e segurança (1), e indefinição e onírico (1); os restantes dois voluntários (16,67%) referem um estado emocional negativo (nostalgia, tristeza) (Lessa, a publicar).

## **5.8 Resultados de utilização do sistema**

Para a avaliação do sistema em tempo-real e com o *wearable*, foi pedido à mesma pessoa que se disponibilizou para a recolha dos sinais fisiológicos que utilizasse o sistema de forma a que se pudesse testar o sistema em condições “reais”.

Verificou-se que com a pessoa sentada e imóvel, o sistema fazia a captação, a análise e a detecção em tempo real dos sinais fisiológicos e comportava-se na globalidade como o esperado: o estado de alegria era detectado, assim como o de repouso e o do



susto. O estado da tristeza não era tão facilmente identificável. Embora tivesse sido detectado com sucesso, não o foi em todas as situações na qual foi pedido à pessoa que pensasse em algo que a pudesse deixar triste. Esta situação pode estar relacionado com duas situações:

- o próprio humor da pessoa na altura em que se realizou a sessão de reconhecimento das emoções em tempo-real;
- a ausência de um estímulo suficientemente forte que a tivesse levado a adquirir um sentimento de tristeza.

Com o *wearable* vestido e com a pessoa em pé e imóvel, verificou-se que a detecção do seu estado emotivo era mais problemática, ou seja, havia estados que não eram detectados correctamente. O estado de repouso consegue ser detectado, assim como o do susto e da alegria, embora tenham sido detectados alguns estados “falsos” que não correspondiam ao estado emotivo da pessoa.

O estado de tristeza é, contudo, o que se revela mais problemático em termos de detecção. Isto deve-se principalmente à questão de que o sinal muscular com o *wearable* vestido e com o monitor TFT ser significativamente diferente do sinal sem o *wearable* e sem o monitor. Uma vez que para segurar o monitor TFT se usam alças sustentadas pelos ombros (tal como uma mochila) e como num dos ombros é colocado um eléctrodo para captação do sinal, esse ombro vai estar sujeito a uma maior actividade muscular, contaminando assim o sinal e impedindo que se faça a detecção correcta dos estados emotivos.

O eléctrodo HR sofre também algumas alterações de sinal quando o monitor TFT é usado, o que se deve ao facto do eléctrodo estar colocado no peito, zona que o TFT pode afectar.

Apesar de na elaboração do fato se ter tido em conta todos estes factores, mostrou-se incontornável alguma alteração dos sinais captados quando este estava vestido.

Outro aspecto no qual o sistema sofre com o ruído dos sinais é em movimento. Uma vez que os dados originais foram todos adquiridos com a pessoa em repouso (apesar de ter sido submetida a estímulos que a levaram a efectuar alguns movimentos, como no caso do susto), o sistema não se comporta correctamente quando a pessoa se encontra em movimento, uma vez que os sinais são muito diferentes dos que foram utilizados para treinar a rede.

## 6 Conclusão

Este capítulo pretende apresentar uma retrospectiva do trabalho realizado, propor melhorias no protótipo e sugerir possíveis desenvolvimentos futuros e, para terminar, são apresentadas considerações finais sobre o sistema desenvolvido e todo o trabalho que lhe esteve subjacente.

### 6.1 Breve retrospectiva do trabalho realizado

O núcleo deste trabalho consistiu no desenvolvimento de um sistema que através da captação dos sinais fisiológicos de um indivíduo (nomeadamente o batimento cardíaco, a sudção cutânea e a actividade muscular) fosse capaz de detectar o seu estado emotivo e afectar em tempo real a produção visual e sonora com base nesses mesmos sinais e estado. Adicionalmente, o sistema deveria analisar o sinal de vídeo obtido através de uma *web cam* e afectar em tempo real a produção sonora do sistema.

Para tal, foi levado a cabo um estudo da Computação Afectiva (computação que se relaciona com, emerge de, ou deliberadamente influencia emoções), desde o seu surgimento até ao seu estado actual. Embora tratando-se ainda de uma área de investigação recente, pode ter um papel importante para o avanço na forma como compreendemos a emoção e a cognição, além de proporcionar a implementação de sistemas informáticos com a capacidade de reconhecer e expressar emoções, desenvolver a sua capacidade de resposta inteligente a emoções humanas, e dar a capacidade a esses sistemas de regular e utilizar as suas emoções.

A análise de aspectos exteriores de um indivíduo e da sua postura ajuda a inferir sobre o seu estado emotivo. Mas esse estado pode também ser inferido, nas emoções mais básicas como a alegria, a tristeza, o susto ou um estado de repouso, através da análise dos seus sinais fisiológicos, nomeadamente através do batimento cardíaco, sudção cutânea e actividade muscular, sendo estes os sinais de base para a detecção do estado emotivo do utilizador.

A análise destes sinais pode ser feita recorrendo a vários métodos, quer através de análise supervisionada, quer não-supervisionada. Na análise supervisionada, uma das técnicas mais utilizadas para o

reconhecimento de padrões numa grande quantidade de dados é a utilização de redes neuronais.

Após esta recolha de informação sobre o actual estado da Computação afectiva, assim como dos métodos de análise necessários para a implementação de um sistema deste tipo, foi feita uma análise das ferramentas que pudessem ser utilizadas para a implementação de um futuro protótipo de Computação Afectiva.

Posteriormente, passou-se para a análise dos requisitos que se pretendia que o protótipo viesse a satisfazer. Foi também definida uma arquitectura geral do sistema em termos de módulos que desempenham funções específicas e que comunicam entre si, sendo todos eles parte integrante do protótipo.

A implementação do protótipo foi um processo demorado que decorreu durante mais de um ano, tendo envolvido vários recursos para a implementação das várias partes constituintes (Módulos) de forma a se conseguir um todo (o protótipo).

O desenvolvimento dos vários Módulos constituintes do sistema foi efectuado de uma gradual e sequencial, uma vez que a implementação de alguns dependia da implementação prévia de outros. Assim, na implementação de qualquer destes Módulos foi dada particular atenção à sua interface com outros Módulos, quer em termos de entrada quer de saída de dados.

Paralelamente, foi desenhado e construído um fato no qual pudesse ser aplicado o protótipo e testado em tempo-real quando a ser utilizado por um indivíduo, a fim de se verificar o seu funcionamento e avaliada a sua resposta.

A avaliação do funcionamento do sistema completo (com o fato, portanto) foi considerada positiva, conseguindo o sistema detectar a maioria dos estados emotivos e respondendo o sistema em termos de imagem e som aos dados fisiológicos do indivíduo. No entanto, o estado emotivo da tristeza não era facilmente detectado, sendo este o estado menos bem conseguido aquando da utilização do sistema.

## **6.2 Desenvolvimentos futuros**

Embora os objectivos propostos para esta Dissertação de Mestrado tenham sido alcançados com sucesso através da implementação do protótipo, este, no entanto, devido a algumas limitações inerentes ao

próprio conceito de um protótipo, não se pode considerar, como é óbvio, como um produto maduro e final. Contudo, mais importante do que as suas limitações actuais são as perspectivas de desenvolvimento futuro que com base no sistema actual podem ser levadas a cabo.

Um dos aspectos de desenvolvimento futuro seria o de dotar o sistema do reconhecimento de um maior número de emoções, além das mais básicas apresentadas neste protótipo (alegria, tristeza, susto e repouso). Adicionalmente, o sistema poderia detectar graus de uma determinada emoção, aumentando ainda mais a capacidade de traduzir fielmente o estado emotivo do utilizador.

Para que tal fosse possível, e para também possibilitar a utilização do sistema por qualquer pessoa (e não apenas por aquela com que foi feita a recolha de dados), teria que se proceder a um maior número de sessões de recolha de dados derivados de estímulos, sessões essas com o maior número de voluntários possível. Deste modo, a rede neuronal poderia ser treinada com um maior número de dados.

Mesmo que não se pretenda detectar um maior número de estados emotivos, ou mesmo detectar graus da emoção em cada estado, um maior número de voluntários para a recolha de dados seria bastante útil para o treino da actual rede neuronal, tornando assim o sistema mais robusto e capaz de detectar com mais exactidão os estados emotivos num maior número de utilizadores (e não apenas nos voluntários das sessões de recolha de dados).

Uma das limitações do protótipo, e como foi dito no Capítulo 5, prende-se com a detecção do estado emotivo da tristeza quando o protótipo se encontra em utilização com o fato vestido, não sendo esse estado tão facilmente detectável como os restantes. Uma possível solução passaria pela elaboração de um fato em que os eléctrodos não estivessem em contacto com este e em que o método utilizado para segurar o monitor TFT ao corpo fosse diferente do actual método das alças sobre os ombros, o que evitaria que os eléctrodos captassem algum ruído causado pelo seu contacto com o fato e evitando também a captação de valores do sinal muscular alterados pelo peso extra do monitor nos músculos dos ombros.

Outro dos aspectos a considerar num desenvolvimento futuro seria o da utilização de um método de análise diferente, em substituição das redes neuronais. Métodos como o SVM (*Support Vector Machine*) e

técnicas de análise de dados que utilizam a projecção de Fisher possibilitam o reconhecimento de emoções com base nos sinais fisiológicos, podendo ser uma alternativa às redes neuronais.

Poder-se-ia também utilizar-se um novo sistema de aquisição de sinais fisiológicos, com um maior número de eléctrodos que pudessem captar outros sinais fisiológicos além do batimento cardíaco, sudação cutânea e actividade muscular. A utilização de um maior número de sinais aliada a outros métodos de análise desses mesmos sinais (ou mesmo através da utilização de redes neuronais adaptadas à entrada de um maior número de sinais fisiológicos) poderiam ajudar à detecção de um maior número de estados emotivos do utilizador.

Em relação à parte visual e sonora do protótipo, poderiam desenvolver-se *templates* visuais e sonoros (ou conjuntos de elementos visuais e sonoros pré-definidos), de forma a que o utilizador pudesse ele próprio escolher o conteúdo visual e sonoro para determinado estado de espírito. Adicionalmente, o utilizador poderia também fazer os seus próprios *templates* de componentes visuais e sonoros, alterando o conteúdo de cada Pacote a seu gosto. Esta funcionalidade obrigaria a uma reestruturação profunda do sistema actual de modo a que se pudesse tornar aberto à inclusão de novos vídeos, imagens e sons. Mas uma vez que neste protótipo se pretendeu apenas mostrar um exemplo de aplicação, optou-se por tornar o sistema fechado e não permitir nesta fase a mudança das componentes visual e sonora por parte do utilizador.

Outros dos aspectos que poderia ser explorado em desenvolvimentos futuros seria o da análise do sinal de vídeo da *web cam* para aumentar a interactividade dos Pacotes, quer em termos visuais quer sonoros. Uma análise mais aprofundada do seu sinal, por exemplo a análise simultânea das componentes de vermelho, azul e verde poderia ser utilizada para a produção musical com diferentes instrumentos associados a cada uma das componentes de cor.

Finalmente, e devido à constante evolução do poder computacional dos equipamentos informáticos, e à evolução em termos de software e hardware de apoio, poderiam ser utilizadas outras ferramentas que complementassem o sistema actual. Por exemplo, foi adicionado recentemente à gama de produtos WaveRider (e após a aquisição do sistema WaveRider Jr. utilizado neste protótipo) um sistema sem

fios (*wireless*), que poderia ser utilizado para dar total mobilidade ao utilizador em relação ao resto de todo o equipamento informático responsável pelo processamento do sistema. Adicionalmente, novas ferramentas de programação, tal como o Max/MSP e Jitter, ou a recente linguagem *processing* baseada em Java, poderiam complementar o sistema em termos visuais e sonoros, oferecendo novos níveis de interactividade e gratificação ao nível da experiência multimédia.

### 6.3 Considerações finais

Os objectivos propostos para esta Dissertação de Mestrado foram alcançadas. Com efeito, foi criado um protótipo capaz de captar os sinais fisiológicos de um indivíduo e utilizar esses sinais para interagir em tempo real com conteúdos multimédia (imagem, vídeo, som e texto). Além disso, e com base nesses mesmos sinais fisiológicos, reconheceu-se o estado emotivo do indivíduo, reflectindo-se esse estado também na interacção com todo o conteúdo do sistema.

O protótipo utiliza o *hardware* WaveRider Jr. e o *software* WaveWare para a captação do sinal do batimento cardíaco (HR), sudação cutânea (GSR) e actividade muscular, sinais estes de base para a detecção do estado emotivo.

A linguagem Pure Data é a responsável por toda a apresentação multimédia do sistema, permitindo que os elementos visuais e sonoros possam ser alterados em tempo-real através da utilização dos sinais fisiológicos.

O Matlab em conjunto com o Java são as ferramentas responsáveis pela detecção do estado emotivo do utilizador com base nos sinais fisiológicos captados. Este estado emotivo condiciona também a apresentação visual e sonora do sistema.

Por fim, todos os estados emotivos evidenciados pelo utilizador, assim como a data e a hora em que ocorreram são guardados pelo sistema numa base de dados de modo a que em qualquer altura se possam consultar os estados já detectados pelo sistema.

Resumindo, o sistema pretendeu ser um espelho das emoções do utilizador. Mas mesmo através da mais complexa análise matemática e do poder computacional das actuais ferramentas informáticas, há

sentimentos tão íntimos que por vezes não se querem transmitir, ficando guardados dentro do utilizador e não sendo detectados fielmente em todas as situações. Poderá assim haver casos em que o que se mostra no exterior através do sistema, não é o que o utilizador realmente sente, podendo assim iludir a máquina, da mesma forma que uma outra pessoa pode ser iludida (há tantos sorrisos carregados de tristeza...).

Por fim, e em última análise, trata-se de preservar a privacidade mais íntima do ser humano, não mostrando tudo de si a qualquer pessoa. Mas apesar dos avanços consideráveis da Computação Afectiva no reconhecimento das emoções, haverá uma parte do ser humano tão íntima e pessoal que apenas o próprio conhece e não mostra, não sendo também detectada? É bom pensar que sim.

## Referências

- Apache; “The Apache Software Foundation”; 2004;  
<http://www.apache.org>
- Bio-Medical; “GSR or Galvanic Skin Response”; 2004;  
<http://www.bio-medical.com/>
- Bio-Medical; “ProComp+”; 2003; <http://www.bio-medical.com/>
- BioControl; “BioMuse”; 2003; <http://www.biocontrol.com/>
- Casanova, Jorge; “Redes Neurais”; 2004;  
<http://www.terravista.pt/aguaalto/1948/RELAT1.html>
- Cycling74; “Jitter”; 2004a; <http://www.cycling74.com/>
- Cycling74; “Max”; 2004b; <http://www.cycling74.com/>
- Cycling74; “MSP”; 2004c; <http://www.cycling74.com/>
- Damásio, António; “Ao Encontro de Espinosa: As emoções sociais e a neurobiologia do sentir”; Publicações Europa-América; 2003
- Day, Scott; “Important Factors in Surface EMG Measurement”, Bortec Biomedical, 2002
- De Luca, Carlo J.; “Surface Electromyography: Detection and Recording”; DelSys Incorporated; 2002
- Dryer, D. Christopher; Ark, Wendy; Lu, David J.; “The Emotion Mouse”; IBM Research Division; 1999
- FM7; “Native Instruments: FM7”; 2004; <http://www.native-instruments.com>
- GEM; “Graphics Environment for Multimedia”; 2003;  
<http://gem.iem.at/>



Gross, James J.; Levenson, Robert W., "Emotion Elicitation Using Films"; Lawrence Erlbaum Associates Limited; 1995

Guidant; "Heart and Blood Vessel Basics "; 2004;  
<http://www.guidant.com/>

IBVA; "Interactive Brainwave Visual Analyzer"; 2003;  
<http://www.ibva.com/>

ICB; "Instituto de Ciências Biológicas"; 2004; <http://www.icb.ufmg.br/>

IEM, "Institute of Electronic Music and Acoustics"; 2003; <http://iem.at/>

Java; "Java Technology"; 2004a; <http://java.sun.com/>

Java; "Java Tutorial: All About Sockets"; 2004b; <http://java.sun.com/>

Java; "Java Tutorial: JDBC Basics"; 2004c; <http://java.sun.com/>

Lessa, Joana; "Fato para Computação Afetiva: protótipo de interface revelador de uma identidade"; Faculdade de Engenharia da Universidade do Porto; a publicar

Loder, Chad; "Neural Networks: An Overview"; 2004;  
<http://www.ccs.neu.edu/home/cloder/>

MathWorks, "Matlab"; 2004a; <http://www.nathworks.com>

MathWorks, "Neural Network Toolbox"; 2004b;  
<http://www.nathworks.com>

MIDIOX; "MIDI-OX: MIDI Yoke Junction"; 2003;  
<http://www.midiox.com/>

MindPeak, "Computerized Instrumentation for Brainwaves and other Biological Signals"; 2003a; <http://www.mindpeak.com/>

MindPeak, "WaveRider Operating Manual"; 2003b;  
<http://www.mindpeak.com/>

MIT; "Affective Computing Group"; 2004a; <http://affect.media.mit.edu/>

MIT; "Affective Computing Projects"; 2004b;  
<http://affect.media.mit.edu/>

MIT; "Research on Sensing Human Affect"; 2004c;  
<http://affect.media.mit.edu/>

Moog, Robert; "Ars Electronica Archive: MIDI - What It Is and What It Means to Electronic Artists"; 2004; <http://www.aec.at/>

Muller, Stefan; "JMatLink"; 2004;  
<http://www.held-mueller.de/JMatLink/>

MySQL; "MySQL: The World's Most Popular Open Source Database"; 2004; <http://www.mysql.com>

Othmer, Siegfried , "Biofeedback - the Ultimate Self-Help Discipline"; 2004; <http://www.eegspectrum.com/>

PD; "Pure Data"; 2003; <http://pd.iem.at/>

PHP, "PHP: Hypertext Preprocessor"; 2004; <http://www.php.net/>

Picard, Rosalind W.; "Affective Computing"; The MIT Press; 1997a

Picard, Rosalind W.; Healey, Jennifer; "Affective Wearables"; Massachusetts Institute of Technology; 1997b

Psychology; "Research on Affective Pattern Recognition and Modeling"; 2004; <http://www.psych.utoronto.ca/>

Pursuit-Performance; "Polar Heart Rate Monitors"; 2004;  
<http://www.pursuit-performance.com.au/>

Rodrigues, Custódio; Teixeira, Marques; Gomes, M. Freitas; "Afectividade"; Contraponto, edições; 1989

Seps, Bert; Beckers, Frank; Ramaekers , Dirk; Aubert, Andre Aubert;  
"Acute heart rate response to weightlessness conditions during  
parabolic flight"; Federal Office for Scientific, Technical and  
Cultural Affairs volume 1, 2001

Shepherd, Peter; "The GSR Meter Course"; Tools for Transformation;  
2001

Uku; "University of Kuopio: Biosignal Analysis and Medical Imaging  
Groups"; 2004; <http://it.uku.fi/biosignal/>